CELL REGULATION, Vol. 2, 135–154, February 1991

Homologous and unique G protein alpha subunits in the nematode *Caenorhabditis elegans*

Michael A. Lochrie,*† Jane E. Mendel,*‡ Paul W. Sternberg,*‡ and Melvin I. Simon*§
‡Howard Hughes Medical Institute
*Division of Biology
California Institute of Technology
Pasadena, California 91125

A cDNA corresponding to a known G protein alpha subunit, the alpha subunit of G_o ($G_o\alpha$), was isolated and sequenced. The predicted amino acid sequence of *C. elegans* $G_o\alpha$ is 80–87% identical to other $G_o\alpha$ sequences. An mRNA that hybridizes to the *C. elegans* $G_o\alpha$ cDNA can be detected on Northern blots. A *C. elegans* protein that crossreacts with an anti-bovine $G_o\alpha$ antibody can be detected on immunoblots. A cosmid clone containing the *C. elegans* $G_o\alpha$ gene (*goa-1*) was isolated and mapped to chromosome I. The genomic fragments of three other *C. elegans* G protein alpha subunit genes (*gpa-1, gpa-2,* and *gpa-3*) have been isolated using the polymerase chain reaction. The corresponding cosmid clones were isolated and mapped to disperse locations on chromosome V. The sequences of two of the genes, *gpa-1* and *gpa-3,* were determined. The predicted amino acid sequences of *gpa-1* and *gpa-3* are only 48% identical to each other. Therefore, they are likely to have distinct functions. In addition they are not homologous enough to G protein alpha subunits in other organisms to be classified. Thus *C. elegans* has G proteins that are identifiable homologues of mammalian G proteins as well as G proteins that appear to be unique to *C. elegans.* Study of identifiable G proteins in *C. elegans* may result in a further understanding of their function in other organisms, whereas study of the novel G proteins may provide an understanding of unique aspects of nematode physiology.

Introduction

G proteins mediate transmembrane signal transduction by physically coupling cell surface receptors to effector proteins that influence intracellular metabolism (Gilman, 1987). A large family of G protein-linked receptors has been identified (Caron, 1989). All of these receptors have seven putative transmembrane segments and bind small organic molecules or peptide hormones. The effectors regulated by G proteins that have been identified include adenylate cyclase, cGMP phosphodiesterase, phospholipases, and various ion channels.

G proteins are composed of three subunits: α, β, and γ. The α subunit undergoes a guanine nucleotide exchange and hydrolysis cycle. In most cells it is the α subunit that directly interacts with effector proteins. G protein α subunits are substrates for several covalent modifications. They are ADP-ribosylated by pertussis toxin, cholera toxin, or both (Moss and Vaughan, 1988). Some G protein α subunits are also myristoylated (Buss *et al.,* 1987). Myristoylation occurs on glycine-2 and is thought to be important for localization of G protein α subunits to the inner face of the cytoplasmic membrane (Jones *et al.,* 1990; Mumby *et al.,* 1990). G protein α subunits are also phosphorylated (Gunderson and Devreotes, 1990; reviewed by Sagi-Eisenberg, 1989).

G proteins are members of a diverse family of proteins with a variety of functions and tissue distributions. G_o, the "other" G protein (Neer *et al.,* 1984; Sternweis and Robishaw, 1984), is abundant in brain and is found only in organisms with nervous systems. G_o has been found in *Drosophila* (deSousa *et al.,* 1989; Schmidt *et al.,* 1989; Thambi *et al.,* 1989; Yoon *et al.,* 1989), rat (Jones and Reed, 1987), cow (Van Meurs *et al.,* 1987), hamster (Hsu *et al.,* 1990), human (Lavu *et al.,* 1988), and *Xenopus* (Olate *et al.,* 1989) but not in yeast or the slime mold *Dictyostelium.* In vertebrates, G_o is concentrated in neuropil (Gabrion *et al.,* 1989; Worley *et al.,* 1986) and growth cones (Strittmatter *et al.,* 1990). In *Drosophila,* G_o has been found in brain and ovaries (deSousa *et al.,* 1989; Schmidt *et al.,* 1989; Thambi *et al.,* 1989; Yoon *et al.,* 1989). G_o may function in several signal transduction pathways (Serventi *et al.,* 1990). In many cases

† Current address: Department of Microbiology and Immunology, University of California, San Francisco, CA 94143.
§ Corresponding author.

it appears to be able to interact with the same receptors as the closely related G_i proteins, although there are some exceptions (Senogles *et al.,* 1990; Ueda *et al.,* 1990). However, G_o may regulate a set of effectors distinct from those regulated by the G_i proteins. The G_i proteins regulate adenylate cyclase and atrial potassium channels, whereas there is evidence that G_o regulates calcium channels (Hescheler *et al.,* 1987; Harris-Warrick *et al.,* 1988; McFadzean *et al.,* 1989), neuronal potassium channels (Van Dongen *et al.,* 1988), and phospholipase C (Moriarty *et al.,* 1989). G_o has no direct effect on adenylate cyclase activity.

To understand the role of G proteins in development and the function of the nervous system and to study G proteins in an organism where an extensive genetic analysis could also be undertaken, we sought to characterize G protein genes in *C. elegans. C. elegans* is a model organism for studying development and the function of the nervous system (Kenyon, 1988; Wood, 1988). It is the only multicellular organism for which a complete cell lineage is known (Sulston, 1988; Sulston *et al.,* 1988). The hermaphrodite contains 959 somatic nuclei; of these, 302 are neuronal. The entire neuroanatomy and neuronal wiring diagram is known (White *et al.,* 1986, 1988; Chalfie and White, 1988). *C. elegans* exhibits many behaviors (reviewed by Chalfie and White, 1988). It has many of the neurotransmitters and signaling systems found in coelomates (Chalfie and White, 1988). Numerous mutants that affect developmental and neuronal processes have been isolated (Hodgkin *et al.,* 1988). Finally, the molecular genetics of *C. elegans* has advanced to the stage where gene transformation is facile (Fire, 1986; C. Mello, V. Ambros, J. Kramer, and D. Stinchcomb, personal communication) and an ordered set of cosmid clones covering most of the genome is available to facilitate gene cloning and mapping (Coulson *et al.,* 1986, 1988).

Our first goal was to determine how many G protein genes exist in *C. elegans.* The characterization of a β subunit gene (*gpb-1;* Van der Voorn *et al.,* 1990) and an α subunit gene (*gpa-2;* Fino Silva and Plasterk, 1990) has been reported. This paper reports the molecular cloning and sequencing of a *C. elegans* $G_o\alpha$ cDNA and the genes encoding two other apparently novel *C. elegans* G protein α subunits (*gpa-1* and *gpa-3*). These cloned genes can provide tools to study their role in *C. elegans* biology.

Results

Polymerase chain reactions

To isolate *C. elegans* G protein α subunit genes, we used short, degenerate oligonucleotide primers in polymerase chain reactions on genomic DNA. The primers correspond to three closely spaced amino acid sequences that are highly conserved in G protein α subunits. They represent three of the six regions that are thought to be involved in guanine nucleotide binding and, therefore, would be expected to be present in most G protein alpha subunit sequences. The distances between the oMP19, oMP20/oMP21 and the oMP19, Tα29 primer pairs in a cDNA would be ~200 and 375 base pairs, respectively. Because genomic DNA was amplified in this study, the PCR products could be larger if they contain introns. PCR amplification of *C. elegans* genomic DNA with the oMP19 and oMP20 primer pair produced PCR products of 270, 525, and 1000 base pairs. Amplification using oMP19 and oMP21 produced PCR products of 260, 300, and 550 base pairs. Amplification using oMP19 and Tα29 produced PCR products of 500, 600, and 800 base pairs. The 550 base pair PCR product from the oMP19 and oMP21 reaction, and the 500, 600, and 800 base pair PCR products from the oMP19 and Tα29 reaction, were subcloned. The sequences of several subclones of each PCR product were analyzed. The sequences fell into three classes designated as *gpa-1, gpa-2,* and *gpa-3.* These sequences contained amino acids that are found in all G protein α subunits (Lochrie and Simon, 1988). The 550 base pair PCR product corresponds to *gpa-2,* the 500 base pair PCR product to *gpa-3,* and the 600 base pair PCR product to *gpa-1.* The 800 base pair PCR product derived by amplification with the oMP19 and Tα29 primers is an artifact of the PCR. Most of the subclones of this PCR product had truncated *gpa-1* or *gpa-3* sequences fused to unknown sequences. Two were *gpa-3/gpa-1* or *gpa-2/gpa-1* hybrid sequences. Out of a total 85 subclones that were sequenced, 86% were *gpa-1, gpa-2,* or *gpa-3* sequences. All three sequences contained introns. The positions of these introns were identical to those of the sixth and seventh introns found in the mammalian $G_i\alpha$, $G_o\alpha$, and $T_r\alpha$ genes (Kaziro *et al.,* 1988; Raport *et al.,* 1989).

$G_o\alpha$ cDNA characterization

A mixture of the three PCR-derived gene fragments was used as a probe to screen a *C. ele-*

gans cDNA library at low stringency. Fifty-two of 45 000 plaques hybridized to the PCR-derived probes. Seven of these clones tested positive by PCR analysis for the presence of G protein α subunit related sequences using the primer pairs oMP19, oMP20/21 or oMP19, Tα29. The insert of one of these clones (λCe6-2) was subcloned. The restriction map and sequencing strategy for this cDNA are shown in Figure 1 and the DNA sequence is shown in Figure 2. The open reading frame found in this sequence encodes a protein that is clearly homologous to $G_o\alpha$ sequences identified in other organisms. The predicted amino acid sequence of *C. elegans* $G_o\alpha$ is 87% identical to *Drosophila* $G_o\alpha$, 80% identical to *Xenopus* $G_o\alpha$, and 81–82% identical to human, rat, mouse, hamster, or bovine $G_o\alpha$ (Figure 3). Two different forms of $G_o\alpha$, called $G_o\alpha$-A and $G_o\alpha$-B, have been described from hamster (Hsu *et al.*, 1990) and mouse (Strathmann *et al.*, 1990). They differ in their sixth and seventh exons as a result of alternative splicing patterns. However, the *C. elegans* $G_o\alpha$ corresponds to neither form of $G_o\alpha$ because it is as different from $G_o\alpha$-A as it is from $G_o\alpha$-B. All of the motifs involved in GTP binding and hydrolysis (Lochrie and Simon, 1988) are highly conserved in the *C. elegans* $G_o\alpha$ sequence. A site for myristoylation on rat $G_o\alpha$, Glycine-2 (Mumby *et al.*, 1990), is present. Amino acids that are substrates for cholera toxin (Arg-179) and pertussis toxin (Cys-351) are also found in *C. elegans* $G_o\alpha$, as they are in other $G_o\alpha$ proteins. The region of *C. elegans* $G_o\alpha$ that is most different from other $G_o\alpha$ sequences is in the region of amino acids 90–140 and 290–320. This divergence would be expected because this region is also the most variable between any other pair of G protein α subunits.

No sequences resembling the SL1 or SL2 spliced leader sequences, which are spliced posttranscriptionally onto certain *C. elegans* mRNAs (Krause and Hirsh, 1987; Huang and Hirsh, 1989), were found at the 5′ end of the $G_o\alpha$ cDNA. However, the $G_o\alpha$ cDNA sequence in Figure 2 is probably not full length, because the size of the mRNA detected on a Northern blot is ~2.2 kilobase pairs (Figure 4). Thus, excluding the poly-A^+ tail, the sequence may be missing ~400 base pairs. The results of Southern blot analysis indicate that the *C. elegans* $G_o\alpha$ gene is present in a single copy in the *C. elegans* genome (data not shown).

To identify a candidate protein encoded by the *C. elegans* $G_o\alpha$ cDNA, we performed an immunoblot analysis with a heterologous antibody. A protein of 40 kDa can be detected on immunoblots of total *C. elegans* extracts with an affinity-purified antibody to bovine $G_o\alpha$ (Figure 5). The intensity of the signal is the same in extracts from the wild-type N2 hermaphrodite strain as in extracts from a *him-5* strain that produces ~30% males. Thus $G_o\alpha$ is probably not male-specific. A protein of the same molecular weight can also be detected with an antipeptide antibody (OC1; McFadzean *et al.*, 1989) that was made against the sequence Ala-Asn-Asn-Leu-Arg-Gly-Cys-Gly-Leu-Phe (data not shown). This sequence is found as the last 10 amino acids of all $G_o\alpha$ proteins except *Xenopus* $G_o\alpha$ and rodent $G_o\alpha$-B.

Isolation and mapping of cosmid clones

The PCR-derived fragments of *gpa-1, gpa-2,* and *gpa-3* and the $G_o\alpha$ cDNA insert were used as

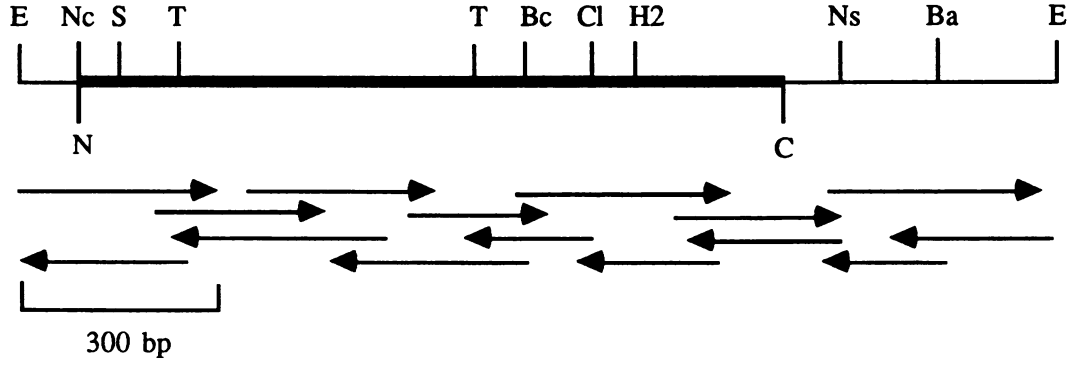

***Figure 1.* Restriction map and DNA sequencing strategy of *C. elegans* $G_o\alpha$ cDNA.** The *C. elegans* $G_o\alpha$ cDNA in λCe6-2 is represented as a horizontal line. The $G_o\alpha$ coding region is represented as a thicker horizontal line. The arrows indicate the direction and extent of DNA sequences determined. Abbreviations for restriction enzyme recognition sites are Ba, *Bam*HI; Bc, *Bcl* I; Cl, *Cla* I; E, *Eco*RI; H2, *Hin*cII; Nc, *Nco* I; Ns, *Nsi* I; S, *Sph* I; T, *Taq* I. Other abbreviations used are: N, amino terminus; C, carboxyl terminus.

```
CGAGCTGCACCACATTACAGTGAGTGAGTAGAGGATATCAAGTGGAGACCGCTACCGGGGCATAGGTCCACCGTT   75
CATCAACTCTAGGTGCCATGGGTTGTACCATGTCACAGGAAGAGCGTGCCGCTCTTGAAAGATCACGAATGATTG  150
                  MetGlyCysThrMetSerGlnGluGluArgAlaAlaLeuGluArgSerArgMetIleG   20

AGAAAAATCTTAAAGAAGACGGCATGCAAGCGGCAAAAGATATCAAACTGCTGCTACTTGGTGCAGGAGAATCAG  225
luLysAsnLeuLysGluAspGlyMetGlnAlaAlaLysAspIleLysLeuLeuLeuLeuGlyAlaGlyGluSerG   45

GAAAATCGACTATTGTAAAACAGATGAAAATTATTCACGAATCAGGATTCACAGCAGAAGACTACAAACAGTACA  300
lyLysSerThrIleValLysGlnMetLysIleIleHisGluSerGlyPheThrAlaGluAspTyrLysGlnTyrL   70

AGCCGGTTGTCTACAGTAACACGGTTCAATCATTGGTCGCTATTTTGCGCGCCATGAGCAACTTAGGCGTTTCAT  375
ysProValValTyrSerAsnThrValGlnSerLeuValAlaIleLeuArgAlaMetSerAsnLeuGlyValSerP   95

TTGGTTCGGCTGACAGAGAGGTAGATGCAAAATTAGTGATGGATGTGGTGGCACGAATGGAGGACACAGAGCCAT  450
heGlySerAlaAspArgGluValAspAlaLysLeuValMetAspValValAlaArgMetGluAspThrGluProP  120

TCTCAGAAGAATTGCTCAGTTCAATGAAACGGTTGTGGGGAGACGCAGGTGTACAGGATTGTTTCTCAAGGAGTA  525
heSerGluGluLeuLeuSerSerMetLysArgLeuTrpGlyAspAlaGlyValGlnAspCysPheSerArgSerA  145

ACGAGTATCAATTGAATGATTCAGCCAAATATTTCCTTGACGACCTGGAAAGGTTAGGAGAGGCAATATATCAAC  600
snGluTyrGlnLeuAsnAspSerAlaLysTyrPheLeuAspAspLeuGluArgLeuGlyGluAlaIleTyrGlnP  170

CAACTGAGCAACATATTCTCCGAACTCGTGTCAAAACAACTGGTATTGTTGAAGTTCACTTCACATTCAAAAATC  675
roThrGluGlnHisIleLeuArgThrArgValLysThrThrGlyIleValGluValHisPheThrPheLysAsnL  195

TCAATTTCAAATTGTTCGATGTGGGAGGTCAAAGATCAGAAAGGAAGAAGTGGATTCATTGTTTCGAAGATGTTA  750
euAsnPheLysLeuPheAspValGlyGlyGlnArgSerGluArgLysLysTrpIleHisCysPheGluAspValT  220

CTGCTATTATTTTCTGTGTTGCCATGTCAGAGTATGATCAACTGTTGCACGAAGATGAGACAACAAACCGAATGC  825
hrAlaIleIlePheCysValAlaMetSerGluTyrAspGlnLeuLeuHisGluAspGluThrThrAsnArgMetH  245

ACGAATCGCTGAAGCTGTTCGATTCGATCTGTAATAACAAATGGTTCACAGATACATCGATTATTCTGTTCCTGA  900
isGluSerLeuLysLeuPheAspSerIleCysAsnAsnLysTrpPheThrAspThrSerIleIleLeuPheLeuA  270

ACAAGAAGGATCTGTTTGAAGAGAAAATCAAGAAAAGCCCGTTAACGATCTGCTTCCCAGAATATTCAGGACGAC  975
snLysLysAspLeuPheGluGluLysIleLysLysSerProLeuThrIleCysPheProGluTyrSerGlyArgG  295

AAGACTACCACGAGGCATCTGCGTATATTCAAGCACAATTTGAGGCTAAAAACAAATCAGCGAATAAAGAAATCT 1050
lnAspTyrHisGluAlaSerAlaTyrIleGlnAlaGlnPheGluAlaLysAsnLysSerAlaAsnLysGluIleT  320

ATTGCCACATGACATGTGCCACAGACACAACTAACATTCAATTTGTGTTTGACGCTGTCACCGATGTGATTATTG 1125
yrCysHisMetThrCysAlaThrAspThrThrAsnIleGlnPheValPheAspAlaValThrAspValIleIleA  345

CCAATAATCTTCGTGGATGCGGCTTGTATTAAGCTGTCGTCTTCGCCGCCCTCTTCTACCATTCTGTGTGTATCT 1200
laAsnAsnLeuArgGlyCysGlyLeuTyrEnd                                              354

TTGCTTACTTTTCCCAAATTTTTAAAGATTCGTTATTTTTCTCATCCCGCGAGTATGCATCTATAATTTGAGAGC 1275
TTTACATGTACATTATGGTTGAACTGTTTTTATTTTTTGGAAAAAGTTGCAATTTCAGATGTATATGGGCTTTTT 1350
TTTAACCTCTGATTCATTCATAATAGTCTCACCGCCCTTCTTTTTCTCAACTAGGATCCCTAATTTTTCTGTCAC 1425
AAAAATAACACATAAAACTCACAATTATTTATTTTTTCCTTTTTTCTTATTTATTATATATTGTCATTATTCACG 1500
CCACATCCCCCGCCCTTCTCACTCCCGTGTAGCCTTATCGCTATTTGAAAACAAGAAAACTTTAAAAATCTAAAT 1575
TCAGTTG                                                                       1582
```

Figure 2. **Sequence of the *C. elegans* $G_o\alpha$ cDNA.** Amino acid abbreviations used are Ala, alanine; Arg, arginine; Asn, asparagine; Asp, aspartic acid; Cys, cysteine; Gln, glutamine; Glu, glutamic acid; Gly, glycine; His, histidine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Phe, phenylalanine; Pro, proline; Ser, serine; Thr, threonine; Trp, tryptophan; Tyr, tyrosine; Val, valine.

probes to isolate the corresponding cosmid clones. The number of cosmids that hybridized to the *gpa-1, gpa-2, gpa-3,* and $G_o\alpha$ probes was three, one, three, and one, respectively. The cosmids that contain the *gpa-1, gpa-2, gpa-3,* and $G_o\alpha$ (*goa-1*) genes mapped to unique po-

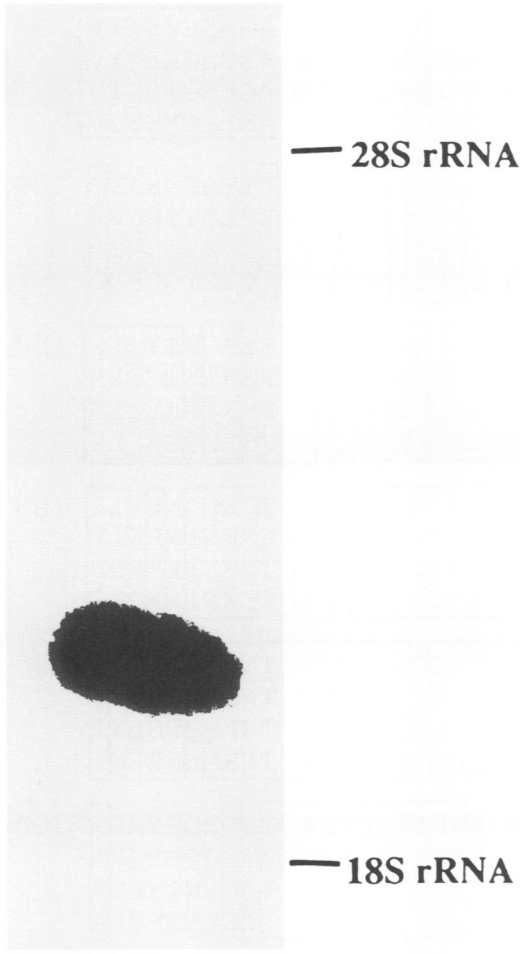

Figure 3. **Hybridization of *C. elegans* total mRNA to *C. elegans* $G_o\alpha$ cDNA.** Total mRNA (12.5 μg) from mixed populations of the N2 wild-type hermaphrodite strain was fractionated on a 1% formaldehyde gel and hybridized to the *C. elegans* $G_o\alpha$ cDNA. The positions of the 18S (1750 bp) and 28S (3500 bp) rRNAs are indicated.

sitions on the *C. elegans* contig map. The *gpa-1, gpa-2,* and *gpa-3* genes map to chromosome V. The *gpa-1* gene maps between the actin (*act-1, act-2,* and *act-3*) gene cluster and *myo-3* and is ~100 kilobases from the actin gene cluster. The *gpa-2* gene maps between *mec-1* and HIS-2, in agreement with Fino Silva and Plasterk (1990). The *gpa-3* gene maps between *hsp-16* and *her-1.* The *goa-1* gene maps to chromosome I between *unc-13* and *lin-10* and is ~50–100 kilobase pairs from *unc-13.* Each gene maps near a number of mutations. Although the positions of many of these mutations on the genetic map are ambiguous, some of these mutations result in phenotypes that suggest they could be in G protein alpha subunit genes.

Gene sequences of gpa-1 *and* gpa-3

The restriction map, intron/exon organization, and DNA sequencing strategy of the *gpa-1* and *gpa-3* genes are shown in Figures 6 and 8. The DNA sequences of the *gpa-1* and *gpa-3* genes are shown in Figures 7 and 9. The *gpa-1* coding region is divided by seven introns and the *gpa-3* coding region is divided by five introns. Intron boundaries were established by several criteria: homology to the amino acid sequences of known G protein α subunit proteins, the spacing between amino acids found in all G protein α subunits, the presence of *C. elegans* consensus splice junction sequences (Emmons, 1988), the absence of stop codons in open reading frames, the absence of any sequence homology to G protein α subunits in putative introns, and conservation of intron position relative to mammalian G protein α subunit genes. In almost every case the 5′ GTAAGTT and 3′ TTTCAG consensus *C. elegans* mRNA splice sites (Emmons, 1988) or closely related variants are found at each intron/exon junction. The sequence GTAAG is found at all of the putative 5′ splice junctions in the *gpa-1* gene and three out of five 5′ junctions in the *gpa-3* gene. The sequence TTXCAG is found at most of the 3′ junctions. An A at position −16 to −17 from the 3′ splice junction border that is conserved—and is believed to be the site where lariat formation occurs during mRNA splicing—is also found in most of the putative introns in the *gpa-1* and *gpa-3* genes, with the exception of some of the smaller ones. Most of the introns are also AT-rich, as observed for other *C. elegans* introns. All of the introns in the *gpa-1* and *gpa-3* genes are >65% AT except for introns 1, 5, and 7 in the *gpa-1* gene. More introns may exist 5′ or 3′ to the coding regions, especially in the 5′ region of the *gpa-3* gene, where several 3′ consensus splice junctions are found (see below).

The sizes and positions of the introns in the *gpa-1, gpa-2,* and *gpa-3* genes are shown in Figure 10. The positions of four introns in the *gpa-1, gpa-2,* and *gpa-3* genes are conserved relative to the mammalian $G_i\alpha$, $G_o\alpha$, and $T_r\alpha$ genes (Itoh *et al.,* 1988; Kaziro *et al.,* 1988; Weinstein *et al.,* 1988; Raport *et al.,* 1989). Those introns in the *gpa-1, gpa-2,* and *gpa-3* genes that occur at the same position as an intron in a mammalian gene not only occur at the same position within the amino acid sequence but also at the same position within the codon. In addition, these introns occur near amino acid sequences that are well conserved (i.e., those involved in guanine

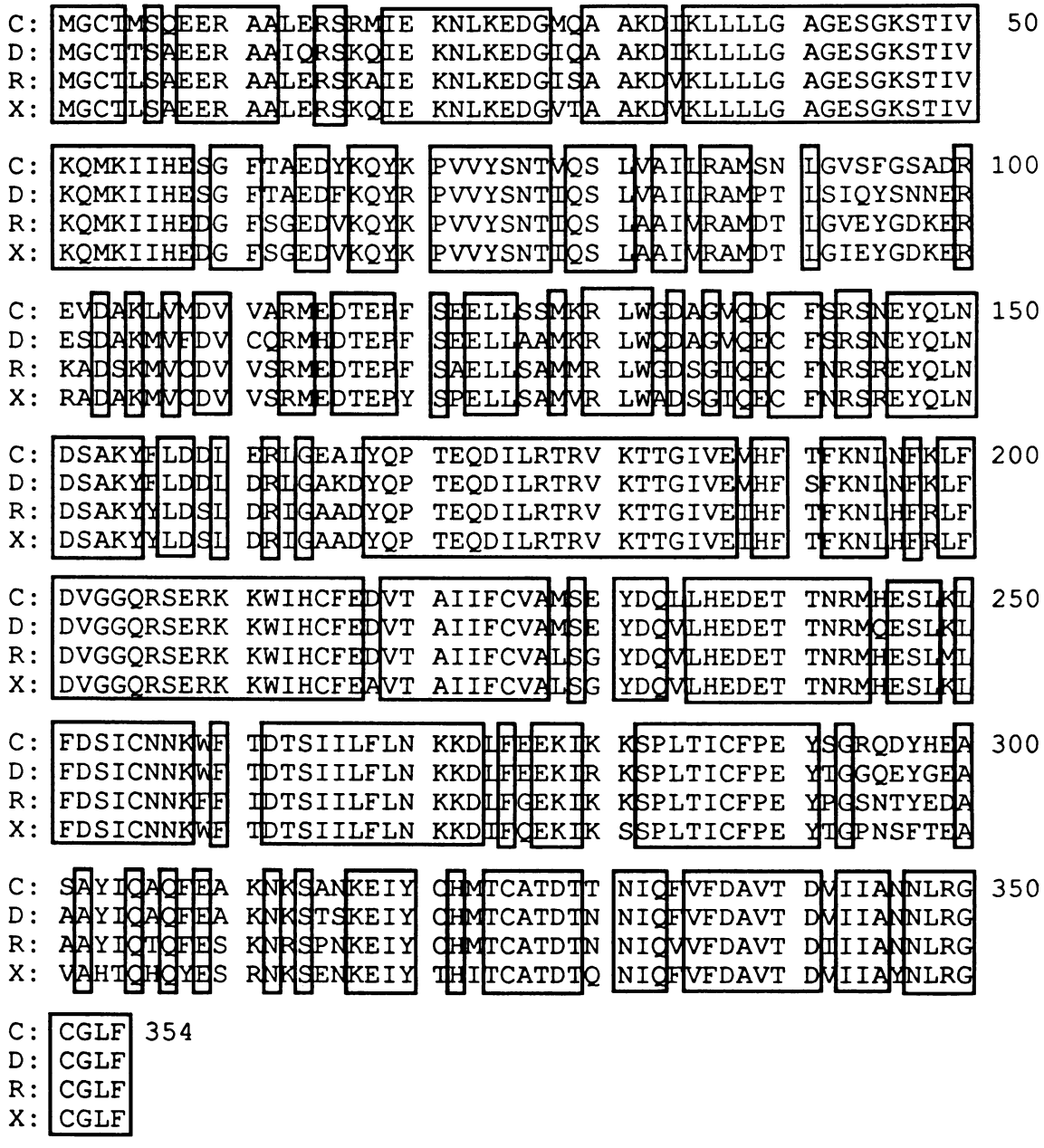

***Figure 4.* Comparison of $G_o\alpha$ amino acid sequences from four species.** $G_o\alpha$ protein sequences from C, *C. elegans* (this study), D, *Drosophila melanogaster* (deSousa *et al.*, 1989), R, rat (Jones and Reed, 1987), and X, *Xenopus laevis* (Olate *et al.*, 1989) were aligned. Amino acids that are found in all four sequences at the same position are boxed. The mouse (Strathmann *et al.*, 1990), human (Lavu *et al.*, 1988), hamster (Hsu *et al.*, 1990), and bovine (Van Meurs *et al.*, 1987) $G_o\alpha$ amino acid sequences are not shown because they are >90% identical to the rat $G_o\alpha$ sequence. Similarly, the locust $G_o\alpha$ sequence (Hsu *et al.*, 1990) is 93% identical to the *Drosophila* sequence. Amino acid abbreviations used are A, alanine; C, cysteine; D, aspartic acid; E, glutamic acid; F, phenylalanine; G, glycine; H, histidine; I, isoleucine; K, lysine; L, leucine; M, methionine; N, asparagine; P, proline; Q, glutamine; R, arginine; S, serine; T, threonine; V, valine; W, tryptophan; Y, tyrosine.

nucleotide binding). The introns in the *gpa-1, gpa-2,* and *gpa-3* genes that are extra or missing relative to the mammalian $G_i\alpha$, $G_o\alpha$, and $T_r\alpha$ genes occur within the region that is most variable between G protein alpha subunits.

A remarkable feature of the intron/exon organization of both genes is that the intron at

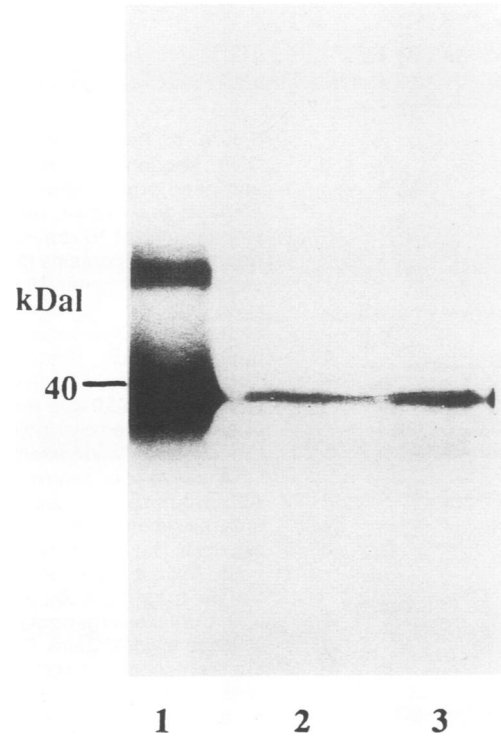

Figure 5. **Immunoblot analysis of *C. elegans* $G_o\alpha$ protein.** An affinity-purified rabbit anti-bovine $G_o\alpha$ antisera was used at a dilution of 1:1000 to detect $G_o\alpha$ proteins as described in Methods. Lane 1, 1 μg purified bovine $G_o\alpha$; Lane 2, 50 μg total protein from *C. elegans* N2; Lane 3, 50 μg total protein from *C. elegans him-5 (e1490).*

the first position within the coding region is unusually large. The first intron of the *gpa-1* gene is 718 base pairs and the first intron of the *gpa-3* gene is 745 base pairs. The first introns of the *gpb-1* gene (1447 base pairs) and the *gpa-2* gene (1020 base pairs) are also large (Fino Silva and Plasterk, 1990; Van der Voorn *et al.,* 1990). This is unusual because in *C. elegans* most of the introns are ~50 base pairs in length (Emmons, 1988). In fact, the average length of the other introns in the *gpa-1, gpa-2, gpa-3,* and *gpb-1* genes is ~130 base pairs. The only extended region of sequence homology that has been detected between any pair of these four large introns is between the first introns of *gpa-2* and *gpa-3.* In this case a stretch of 85 base pairs with 63% nucleotide sequence identity was found. No open reading frames resembling G protein α subunit sequences have been detected within them, indicating that they do not contain alternatively spliced exons. At the other extreme, the third intron in the *gpa-1* gene, which is 43 base pairs long, is among the smallest introns known.

Another interesting feature of the *gpa-1* gene is the presence of an almost perfect inverted repeat (see Figure 6C). One repeat element is located at positions 583–918 within the first intron and the other is at positions 2012–2350 within the fifth intron. The elements of the repeat are ~95% identical. The repeats are ~330 base pairs long and are ~1100 base pairs apart. This corresponds well to the average dimensions of repeats in *C. elegans* DNA reported by Emmons *et al.* (1980). Palindromic sequences are found at the boundaries of the repeats.

The nucleotide sequences 5′ to the initiation codons of the *gpa-1* and *gpa-3* genes were examined for the presence of consensus CAAT-box, TATA-box, and transcription initiation site sequences (Breathnach and Chambon, 1981). The *gpa-1* gene has sequences at positions 50–55 (GCTAAT), 86–92 (TATATA), and 113 that conform well to the spacing constraints and consensus sequences for the CAAT-box, TATA-box, and transcription initiation site. In addition, a consensus 3′ splice junction acceptor site (TTTCAG) is found immediately preceeding the initiation codon. This site may be where a *trans*-splicing reaction occurs. A very similar pattern of consensus CAAT-box, TATA-box, transcription initiation site, and 3′ splice junction sequence is found in the 5′ region of the *gpa-2* gene (Fino Silva and Plasterk, 1990).

The 760 base pairs of sequence that is available 5′ to the *gpa-3* coding region has only one TATA sequence located at position 721–724, which is 35 base pairs from the translation initiation codon. No sequences resembling the CAAT-box sequence are readily found. However, the region from positions 6 to 52 has 68% nucleotide sequence identity to a region in the *gpa-2* gene that is between its putative CAAT-box and TATA-box. Based on this homology, positions 3–7 (GCATT) and 38–43 (TTTATG) may be the CAAT-box and TATA-box sequences of the *gpa-3* gene, respectively. Several potential 3′ splice junctions are found at positions 83–90 (TTTTCCAG), 253–257 (TTTAG), 320–326 (TTTTCAG), 694–700 (TTTTTAG), and 730–736 (TTACAG). However, no sequences resembling the consensus 5′ splice junction sequence are found within this region.

The sequences 5′ to the initiation codon were compared with each other to detect sequence homologies that might, for example, be indicative of similarities in gene regulation. Other than the homology between *gpa-2* and *gpa-3* mentioned above, pairwise sequence comparisons of available sequences from the 5′ regions

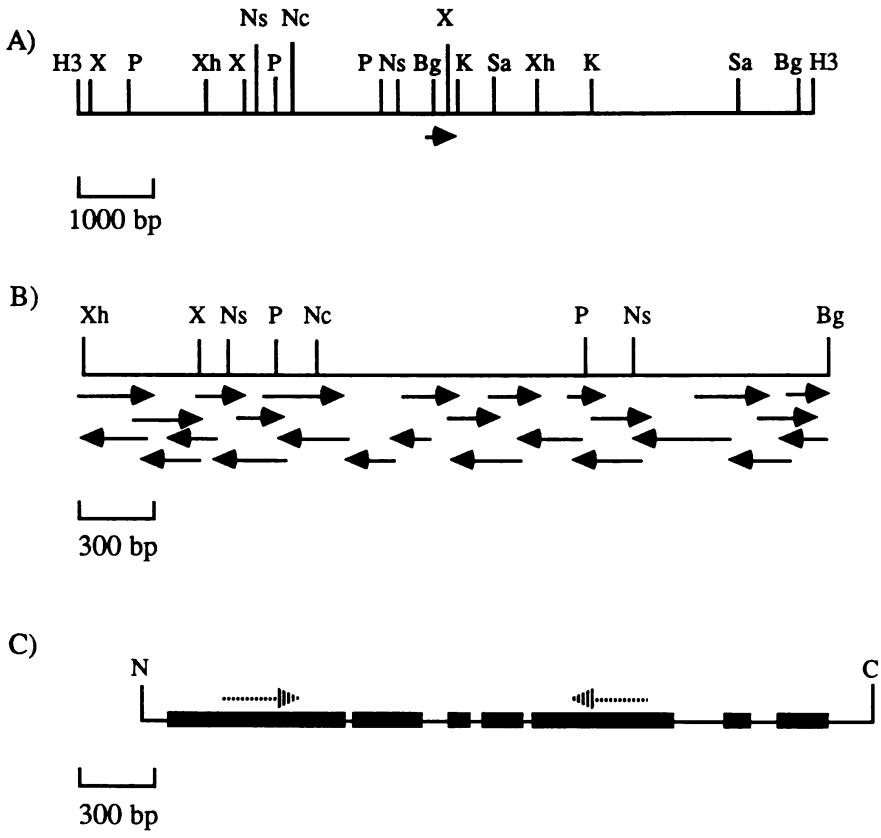

***Figure 6.* Restriction map, DNA sequencing strategy, and gene organization of *C. elegans gpa-1* gene.** (A) Restriction map of 10 kbp *Hin*dIII fragment that contains the *C. elegans gpa-1* gene. The arrows indicate the direction and extent of DNA sequences determined. (B) Restriction map of 3.1 kbp *Xho* I/*Bgl* II subclone from 10 kbp *Hin*dIII fragment. The arrows indicate the direction and extent of DNA sequences determined. Abbreviations for restriction enzyme recognition sites are Bg, *Bgl* II; H3, *Hin*dIII; Nc, *Nco* I; Ns, *Nsi* I; K, *Kpn* I; P, *Pst* I; Sa, *Sal* I; X, *Xba* I; Xh, *Xho* I. (C) Intron/exon organization of *C. elegans gpa-1* gene. Exons are represented as horizontal lines. Introns are represented as boxes. Striped arrows indicate the location of the inverted repeats. Abbreviations used are N, amino terminus; C, carboxyl terminus.

of the *gpa-1, gpa-2, gpa-3,* and *gpb-1* genes did not reveal any striking homologies. Each 5′ region has several repeated sequences, but none of these are common to any pair of sequences. The open reading frames on both strands that are found 5′ and 3′ to the coding regions of the *gpa-1* and *gpa-3* genes were compared with the GenBank database, but no significant homologies were detected.

The available sequences 3′ to the termination codons of the *gpa-1* and *gpa-3* genes were examined for an AATAAA consensus polyadenylation signal sequence (Proudfoot and Brownlee, 1979; Wickens and Stephenson, 1984). This sequence was found in the 3′ region of the *gpa-1* gene at positions 3340–3345. It was not found in the *gpa-3* 3′ region, although minor variants that might be functional are found at positions 2825–2830 (AATACA) and 2891–2896 (AATATA).

The predicted amino acid sequences of the *gpa-1, gpa-2,* and *gpa-3* proteins are compared in Figure 11. The percent sequence identity observed between them is *gpa-1* versus *gpa-2:* 41%, *gpa-1* versus *gpa-3:* 48%, and *gpa-2* versus *gpa-3:* 58%. The *gpa-1, gpa-2,* and *gpa-3* proteins are 49%, 46%, and 53% identical to the *C. elegans* $G_o\alpha$ sequence, respectively. All three *gpa* proteins have the six regions implicated in binding guanine nucleotides (Lochrie and Simon, 1988). Also, all three have the consensus myristoylation signal Met-Gly-X-X-X-Ser at the amino terminus (Towler *et al.,* 1988). Moreover, all three have the arginine (Arg-179 in *gpa-1,* Arg-181 in *gpa-2,* and Arg-179 in *gpa-3*) that is the substrate amino acid for cholera toxin in other G protein α subunits. The *gpa-1* and *gpa-3* proteins have the cysteine (Cys-354 in *gpa-1* and Cys-351 in *gpa-3*) four amino acids from the carboxyl terminus that is the substrate amino acid for pertussis toxin in other G protein α subunits. However, the *gpa-2* protein has a serine at this position (Fino Silva and Plasterk, 1990). The presence of cholera and pertussis toxin substrates in *C. elegans* has been reported (Van der Voorn *et al.,* 1990) but the effects of these toxins on *C. elegans* physiology has not been well studied.

Discussion

We have determined the sequence of a cDNA encoding a *C. elegans* $G_o\alpha$ protein that is very

similar to that of $G_o\alpha$ proteins found in other species. Therefore, it is likely that *C. elegans* $G_o\alpha$ is similar in function to other $G_o\alpha$ proteins. In vertebrates and *Drosophila,* $G_o\alpha$ is found predominantly in neurons. However, its role in signal transduction pathways is not well delineated. Thus, the presence of a $G_o\alpha$ homologue in *C. elegans* may provide new approaches for understanding its function in other organisms.

The results of Southern blot analysis indicate the *C. elegans goa-1* gene is single copy, and only one size class of mRNA is observed on a Northern blot using the *C. elegans* $G_o\alpha$ cDNA as a probe. It is, therefore, likely that the protein detected on the immunoblot is the same as that encoded by the $G_o\alpha$ cDNA described here and not some other closely related G protein α subunit. However, the possibility that there may be multiple $G_o\alpha$ proteins in *C. elegans* cannot be excluded until the genomic sequence of *goa-1* has been determined. Multiple species of $G_o\alpha$, which arise as a result of alternative mRNA splicing patterns, have been found in *Drosophila* (deSousa *et al.,* 1989; Thambi *et al.,* 1989; Yoon *et al.,* 1989) and rodents (Hsu *et al.,* 1990; Strathmann *et al.,* 1990).

In contrast to *C. elegans* $G_o\alpha$, the *gpa-1, gpa-2,* and *gpa-3* proteins cannot be classified according to their relatedness to other G protein α subunits. They are 40–50% identical to any other non-*C. elegans* G protein α subunit sequence. G protein α subunits that are >80% identical are generally considered to be in the same class (Lochrie and Simon, 1988) and can activate the same effector proteins (Gillespie and Beavo, 1988; Yatani *et al.,* 1988). G protein α subunits that are as little as 60–70% identical (e.g., $G_o\alpha$, $G_i\alpha$, and $T_r\alpha$) can interact with some of the same receptors, although heterologous interactions are somewhat less efficient than homologous ones (Kahano *et al.,* 1984; Cerione *et al.,* 1986). However, G proteins in this class have not been found to regulate the same effector proteins. G protein α subunits that are 40–60% identical have not been observed to interact with the same receptor or effector and are generally considered to be distinct in function. Therefore, *gpa-1, gpa-2,* and *gpa-3* are likely to have separate functions. This divergence does not exclude the possibility that they may operate in signaling systems similar to those in mammals, because the receptors and effectors that *gpa-1, gpa-2,* and *gpa-3* interact with may also be different in sequence from their mammalian counterparts.

Although the sequences of the *gpa-1, gpa-2,* and *gpa-3* proteins provide no clues about their function, several approaches are available in *C. elegans* for investigating their function. One approach initiated in this study is the mapping of genes near known mutations with visible phenotypes, followed by a determination of which of these mutations, if any, correspond to mutations in G protein alpha subunit genes. This determination can be done by transformation of mutant strains with the cosmid clones isolated in the work described here. However, it is possible that none of the existing mutations will be in a G protein gene, because the genetic map is not saturated with mutations. If so, there are a number of strategies to identify mutant phenotypes for these genes. For example, recessive mutations might be isolated by a PCR-based screen for transposon insertion mutations, using methods similar to those developed in *Drosophila* (Ballinger and Benzer, 1989; Kaiser and Goodwin, 1990). Dominant mutations can be generated by transformation with mutants of the *gpa* genes constructed *in vitro.* Based on studies of mutant G protein α subunits found in human tumor cells (Landis *et al.,* 1989) and the biochemical properties of G protein α subunits synthesized in *E. coli* (Freissmuth and Gilman, 1989; Graziano and Gilman, 1989) and mammalian cells (Masters *et al.,* 1989), certain amino acid changes in the *C. elegans* G protein alpha subunit sequences might be expected to result in a dominant phenotype. A determination of *gpa* gene expression patterns could also provide information about the function of the *gpa* proteins. However, our initial attempts to detect *gpa* gene expression by Northern blotting were unsuccessful. In addition, we were unable to isolate cDNA clones for *gpa-1, gpa-2,* or *gpa-3.*

gpa-2 and *gpa-3* are the only two known G protein α subunits that have a cysteine instead of a serine in the highly conserved sequence GAGESGK found at the amino terminus (Results; Fino Silva and Plasterk, 1990). The sequence of $G_z\alpha$ also differs in this region from other G protein α subunits. It has the sequence GTSNSGK. These differences found in $G_z\alpha$ are thought to be responsible for its very low GTPase activity (Casey *et al.,* 1990). This region is close to the phosphates of GDP in the crystal structure of the analogous *ras* proteins (Milburn *et al.,* 1990) and mutations in this region of $G_s\alpha$ reduce its GTPase activity (Graziano and Gilman, 1989). Thus, *gpa-2* and *gpa-3* may have a GTPase activity that differs from *gpa-1.* Such

```
CTCGAGTTGTTTATGCGGCGAAAAAAAATCTGTAGAGGTATACGAAAAAGCTAATTAGGT   60
GGTTTGCCGTCTAGTAATACTCGATTATATATTTCTTGTCGTCATCGTTACCACTGCTCT  120
CCATTTTCTCTCCTCACTCACAAAAACTTGAATATTCTTCGCTTCCTGTTGTGATCAGTG  180
AAAATTCGTCTGATTTCTTCAATATTATTCAATGTCATTTATTCGAAGACTTCAGATGGG  240
                                                        MetGl

AAACTGCGAAAGTCGAGAACTTGTAGCTCAAGCCAAACAGAATAAAATAATCAACACGGA  300
yAsnCysGluSerArgGluLeuValAlaGlnAlaLysGlnAsnLysIleIleAsnThrGl

ACTTGATAAGGCCAAAAAAACCGATGAAAATATAATCAAATTATTGCTGTTAGGTAAGTT  360
uLeuAspLysAlaLysLysThrAspGluAsnIleIleLysLeuLeuLeuG

AGATTCTGAACCAAAAACAGCTTTCTGCAATACTTGTTGGTACATTCAAATCGCGCTCAT  420
TTTTCATTGAAAATTAAACGCGGCGAACATTTTTCTTACAAATAGGTGGTGACAACCTCA  480
TTTCATTCTGGTATCGGCACGAAAACATGTTTCTAGAAAGGTCTAGTTACCCAACCAAAA  540
AGTTGAAAGCACCACAGCTGTAGCGAAACCCAAAATAGAAGGAATTAAAATCTACCCACT  600
TGAGTTCACTAACACCCGCAAATGCATTAAAGAAACGTTTAAAAAACATCAGGTTGAAAT  660
CTTGTTCAATTATATCCGCTCAGAAGAGCAAGCCACTACGGCTCCAAAATTCATCGGCTC  720
CGTGCGGTCATCAAGTTCCGAAGAACTGTTTCACGCTCGCGGCTCGCCCGTCTCCCACTC  780
ATCGCGGATGGCATCTCTACCAATTTGATGGGAATTACGAGATACATACTGCAGAAATGA  840
TTCTGTATAGTATGGTCTCGATGTAAGCAGAGTGCTGACGGGGTTAGTGTGTACACGACA  900
GCCGACACCTCGCGGGTTTAGCCTGCCACCCTGTTTTGTGTGTCTGCCTGACCACCTGCC  960
GCCGATTGCATTGATAGCATTGATGCTCCATGGAAATGCGATTACTAATCAAACATCATT 1020
CGAATGTTTGACTTTTCTATCGTTTAAACATCTGAAATTAACTATTTCCAGGAGCTGGAG 1080
                                                    lyAlaGlyG

AGAGTGGAAAAAGTACCGTGCTGAAACAAATGAAGTAAGTTCTTAAAGATGAGTTCGAAA 1140
luSerGlyLysSerThrValLeuLysGlnMetLy

AAGAGTGTTTGCTATAAAATGCTCGGAATAATGCAATATGGGGCAAATGACGGATTAATC 1200
TTTCAGCTTTTATTTTAAGAGGAAGTTCAAAAATTTACAATTATTATTAGATTTTTTTTC 1260
AAAATTTCTTTTTTGTGGCATTTCATCATATTGGAATTTTCCAAGGTCTTTAAAGCAAAA 1320
GCTAAGTTAACCGAACTCTACTTTCAATCAAATGAAGAAATGTACCAATAAAATGATTTA 1380
CAGAATCATCCATAACAGTGGATTCTCTCAAGAAGAAATATCAAATAAACGAAATGTTGT 1440
   sIleIleHisAsnSerGlyPheSerGlnGluGluIleSerAsnLysArgAsnValVa

CTGTGCGAATACTGTGCAAGCAATGGGAGCATTGTTAGATGGAATGAAACAACTTCAATT 1500
lCysAlaAsnThrValGlnAlaMetGlyAlaLeuLeuAspGlyMetLysGlnLeuGlnPh

CGATTTTTCAACCCGAGTTTGTAATGTAAGAAATGTTAATTTTTGTGTTACATAGTACTT 1560
eAspPheSerThrArgValCysAsn

ATTTATAGGCACATGAGAAGTTGATACGTGAAACATTGAATGATAAAGCTGAGTATGGAC 1620
        AlaHisGluLysLeuIleArgGluThrLeuAsnAspLysAlaGluTyrGlyP

CATTCAGTGATGCAATGTTCAAGTAAGAACCTTATGTCATTTTGTGAATTTATTCAAAAA 1680
roPheSerAspAlaMetPheAs

ATATTTTGATAGCAATTCTCCGATTTGACAAAACTTTCATAACAATGGAAGTTCACAATA 1740
AAATTTTAATAAATCTAAACTATATTCAGCGCACTTACTGAGTTGTGGGCGGACAAAGGA 1800
                                nAlaLeuThrGluLeuTrpAlaAspLysGly

GTTCAGTGTGCATACGATAAGCGAGAGTTTTTTTACCTTCATGATTCTGCAAAATAGTAA 1860
ValGlnCysAlaTyrAspLysArgGluPhePheTyrLeuHisAspSerAlaLysTy
```

Figure 7. **DNA sequence of *C. elegans gpa-1* gene.** The predicted amino acid sequence of the *gpa-1* protein is shown below the DNA sequence. The standard three-letter amino acid abbreviations are given in the legend to Figure 2. Coding sequences are shown in boldface lettering. The primers used in PCR reactions are underlined. oMP19 corresponds to nucleotide positions 2564–2581, oMP20/oMP21 to positions 2791–2808, and Tα29 to positions 3140–3157.

```
GTTAAACTTTTTTTTAAATATCAGAACTTTGCGATATTTTAAGTATTTAAAAACATAGAG 1920
CGTTAGAGGCTCAATATTAGTTCAACAAAACAGGTTATTCCATTTCCTAATTGCAATCAA 1980
GAGGCGTGTCTGTGCCTACACTTAGGATGCTGAATGACTGAACCCGCGAGGTGTCGGCTG 2040
TCGTGTACACACTAGCCCCGCCAGCACTCTGCGTACATCGAGACCATACTATACAGAATC 2100
ATTTCTGCAGTATGTATCTCGTAATTCCCATCAAATTGGTAGAGATGCCATCCGCGATGA 2160
GTGGGAGACGGGCGAGCCGCGAGCGTGAAACAGTTCTTCGGAACTTGATGACCCGCACGG 2220
AGCCGATGAATTTTGGAGCCGTAGTGGCTTGCTCTTCTGAGATCGGATATAATCGAACAA 2280
GGTTTCAATCTGATGTTTTTTTAAACGTTTTTTTAATGCATTTGCGGGTGTCTAGTGCGA 2340
TTTTATAATTGAATTACATTAAAAATTGCGAGTAGGTAAATAAGATTTAATTTTCAGTTT 2400
                                                          rPh

TCTTGATCGAATTGCAAGGGTTCACACTCCCAATTACGTACCAACTGAAAATGATATTCT 2460
eLeuAspArgIleAlaArgValHisThrProAsnTyrValProThrGluAsnAspIleLe

GCATACACGGGTTCCAACAATGGGTGTTATAGAAGTTAATTTCACAATAAAGGGCAAATT 2520
uHisThrArgValProThrMetGlyValIleGluValAsnPheThrIleLysGlyLysPh

TTTTCGAGTATTTGATGTGGGCGGACAGCGATCGCAAAGGAAGAAATGGATTCACTGTTT 2580
ePheArgValPheAspValGlyGlyGlnArgSerGlnArgLysLysTrpIleHisCysPh

TGATGATGCAAAAGCTATGATATATGTTGCTTCGTTGTCAGAGTATGATCAAGTTTTATT 2640
eAspAspAlaLysAlaMetIleTyrValAlaSerLeuSerGluTyrAspGlnValLeuLe

AGAAGACAATACTACTGTAAGTTTTGATATTTTCAAAAACAAATAATTCAATAGTAATTT 2700
uGluAspAsnThrThr

CTTGTTTTTCAGAATCGAATGCACGAATCAATACAGTTATTTAAGCAAGTAATCAACAAT 2760
           AsnArgMetHisGluSerIleGlnLeuPheLysGlnValIleAsnAsn

AAATATTTTGTAAACACTTCAGTTATATTATTCCTGAACAAAATTGATTTATTCGAGGAA 2820
LysTyrPheValAsnThrSerValIleLeuPheLeuAsnLysIleAspLeuPheGluGlu

AAAATTGTTACCAAAAAACGAAGTCTTGGTATTGCATTTGAGTCATTTAGTGGTAAGTGC 2880
LysIleValThrLysLysArgSerLeuGlyIleAlaPheGluSerPheSerG

AATTTTTGAAGGCGAATGTGGAGGCACTTGAACCCCGTACACCTGTCCGCCGTTCTGGCC 2940
GAGTGGTCTAAGGCGCTGCGTTCAGGTCGCAGTCCTCTCCGGAGGGCGCAGGTTCGAATC 3000
CTGCGGACGGCATTTCTTTTTGTGCATTCTCTTTTTTTTTGCAGGACCGAGCCAAGATCT 3060
                                                  lyProSerGlnAspLe

CAATGCAGCTGTAGCGTTCGTTGAAAAAAAGTATAGAAGTATGGCAGAGAACAAAGAGAA 3120
uAsnAlaAlaValAlaPheValGluLysLysTyrArgSerMetAlaGluAsnLysGluLy

GAACATTTATTGTCATCACACTTGTGCTACAGACACACAACAGGTGCAGTACGTACTCGA 3180
sAsnIleTyrCysHisHisThrCysAlaThrAspThrGlnGlnValGlnTyrValLeuAs

TGCGGTTCTAGATACGATACTATCTACCAAACTGAAGGGATGTGGATTGTATTGAGAGAT 3240
pAlaValLeuAspThrIleLeuSerThrLysLeuLysGlyCysGlyLeuTyrEnd

TTTAAATTATTGTTACCTTTGGTACAAACCTTGTAAACCTGAGTCATAATTATCGGGTAC 3300
CGTTCGGTATTATCTTGATTGGTGTAATATATAGAATTTAATAAATAATGAATCAAATTT 3360
```

***Figure 7.* (Continued)**

differences could have a profound effect on their biological properties.

There may be G protein genes in *C. elegans* other than those reported to date. A *C. elegans* homologue of the mammalian $G_s\alpha$ gene (*gsa-1*) has been isolated (S. Ohshima, S. Tashiro, J.-H. Park, T. Tani, and Y. Ohshima, personal communication). This gene could not have been iso-

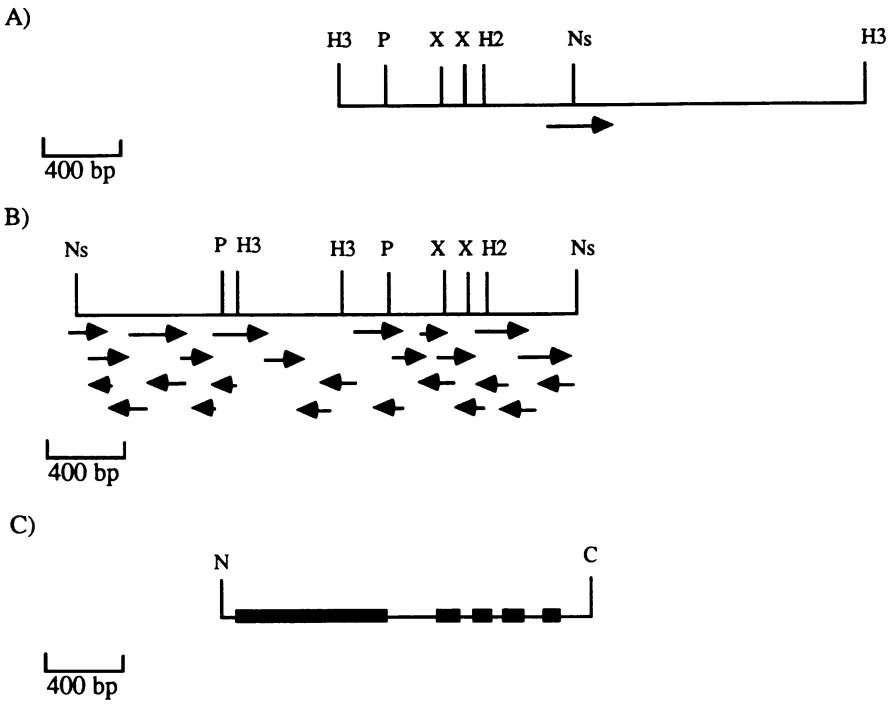

Figure 8. **Restriction map, DNA sequencing strategy, and gene organization of *C. elegans gpa-3* gene.** (A) Restriction map of 3 kbp *Hin*dIII fragment from *C. elegans gpa-3* gene. The arrows indicate the direction and extent of DNA sequences determined. (B) Restriction map of 2.7 kbp *Nsi* I fragment from *C. elegans gpa-3* gene. The arrows indicate the direction and extent of DNA sequences determined. Abbreviations for restriction enzyme recognition sites are Ns, *Nsi* I; H2, *Hin*cII; H3, *Hin*dIII; P, *Pst* I; X, *Xba* I. (C) Gene organization of *C. elegans gpa-3* gene. Exons are represented as horizontal lines. Introns are represented as boxes. Abbreviations used are N, amino terminus; C, carboxyl terminus.

lated by the methods used in this paper because $G_s\alpha$ is too divergent from other G proteins. Furthermore, the PCR analysis reported here suggests that other G protein alpha subunit genes may exist, because not all of the PCR products can be accounted for by the clones that are available. It is also possible that a given size class of PCR product may contain more than one G protein α subunit sequence, as observed by Strathmann *et al.,* (1989). *C. elegans* may have other G proteins that have so far been found only in mammals, such as $G_z\alpha$ and $G\alpha 11$–16 (M. Strathmann, T. Wilke, and T. Amatruda, personal communication), or G proteins that have been found in both *Drosophila* and mammals, such as $G_i\alpha$ (Provost *et al.,* 1988) and $G_q\alpha$ (Strathmann and Simon, 1990). Conversely, other organisms may have homologues of the *C. elegans gpa-1, gpa-2,* and *gpa-3* genes. Finally, because G proteins are always found as heterotrimers, it is likely that *C. elegans* also has G protein γ subunit genes.

In conclusion, *C. elegans* has at least five G proteins. Two of these (G_o and G_s) are identifiable homologues of mammalian G proteins, whereas three (*gpa-1, gpa-2,* and *gpa-3*) appear to be unique to *C. elegans.* All multicellular organisms that have been examined have G_o and G_s. However, unicellular organisms, such as yeast and *Dictyostelium,* do not have G_o and G_s. Instead they have G proteins with novel sequences. Therefore, it may be that G proteins such as G_o and G_s arose during the development of multicellular organisms and have been highly conserved across phylogeny because they have fundamental roles in intercellular signaling. On the other hand, the G proteins such as *gpa-1, gpa-2,* and *gpa-3* that are not well conserved may have functions specific to a particular organism.

Methods

Nematodes

The wild-type hermaphrodite N2 (Bristol) strain is from Brenner (1974) and *him-5 (e1490)* is from Hodgkin *et al.* (1979).

Polymerase chain reactions (PCRs)

Twenty-five nanograms of *C. elegans* genomic DNA from the hermaphrodite wild-type N2 strain were amplified in a 10-μl reaction containing 50 mM KCl, 1.5 mM $MgCl_2$, 10 mM tris(hydroxymethyl)aminomethane (Tris)-HCl, pH 8.3, 0.01% gelatin, 200 μm deoxyribonucleotides (dATP, dCTP, dGTP, and dTTP), 2.5 U Taq polymerase (Perkin Elmer-Cetus, Norwalk, CT), and 0.3 μg of each primer. The primer pairs that were used are oMP19, oMP20/oMP21 and oMP19, Tα29. The primers oMP19, oMP20, and oMP21 have been described (Strathmann *et al.,* 1989). The primer oMP19 corresponds to the sense strand that encodes the amino acid sequence Lys-Trp-Ile-His-Cys-Phe/Leu. The primers oMP20 and oMP21 correspond to the antisense

strand that encodes the amino acid sequence Phe-Leu-Asn-Lys-Lys-Asp. The sequence of Tα29 is GAATTC(GATC)-GT(AG)TC(GATC)GT(GATC)GC(AG)CA(GATC)GT. This sequence contains an *Eco*RI restriction site at one end to facilitate subcloning and corresponds to the antisense strand that encodes the amino acid sequence Thr-Cys-Ala-Thr-Asp-Thr. Thirty amplification cycles consisting of 1 min at 92°C, 30 s at 37°C, and 1 min at 72°C were performed. At the end of the amplification cycles, the reactions were incubated another 10 min at 72°C to complete partially extended chains. The products of the reaction were analyzed on 3% NuSieve (FMC, Rockland, ME) agarose gels. The PCR products were excised from the gel, reamplified in a 100-μl reaction using the same primers, made blunt with Klenow, phosphorylated with T4 polynucleotide kinase, ligated to linkers, and subcloned (Maniatis *et al.,* 1982) into pBluescript KS/− (Stratagene, Burlingame, CA). The DNA sequences of the subclones were determined as described below.

Isolation of $G_o\alpha$ cDNA

A *C. elegans* cDNA library containing ~250 000 clones was constructed in λgt10 and kindly supplied by Stuart Kim (Stanford Univ.). The mRNA used to construct this library represented all developmental stages of the hermaphrodite wild-type N2 strain. The library was plated on 15-cm LB agar plates and replicated to nitrocellulose filters (Maniatis *et al.,* 1982). The filters were prehybridized in 6× SSC (1× SSC is 0.15 M NaCl, 0.015 M sodium citrate, pH 7.0), 0.1% sodium dodecyl sulfate (SDS), 30% formamide, 100 μg/ml sonicated salmon sperm DNA, 5× Denhardt's solution (1× Denhardt's solution is 0.02% Ficoll 400, 0.02% bovine serum albumin (BSA), 0.02% polyvinylpyrrolidone) at 37°C for 2 h. Probes were added to the prehybridization solution and hybridized under the same conditions for 24 h. The probes consisted of the PCR-derived genomic fragments of the *gpa-1, gpa-2,* and *gpa-3* genes. These were isotopically labeled with a Multiprime kit (Amersham, Arlington Heights, IL) using α-^{32}P-dATP. The probes were denatured by boiling for 5 min before use, and 1 × 10^6 cpm/ml of each probe was added to the hybridization solution. After hybridization, the filters were washed with 0.1× SSC, 0.1% SDS at 23°C and exposed to X-ray film for 48 h.

Northern blot analysis

Total RNA was isolated by the guanidine isothiocyanate procedure described by Maniatis *et al.* (1982) from the N2 strain. The RNA was denatured, fractionated on a 1% formaldehyde gel, and blotted to a nitrocellulose filter. The insert from the $G_o\alpha$ cDNA was isotopically labeled using α-^{32}P-dCTP by the method of Feinberg and Vogelstein (1983). The filter was hybridized to the probe in 50% formamide, 5× SSC, 5× Denhardt's solution, 25 mM $NaPO_4$, 0.1% SDS, and 0.25 mg/ml salmon sperm DNA at 42°C. The filter was washed in 0.2× SSC, 0.1% SDS at 60°C and exposed to X-ray film.

Immunoblot analysis

Mixed populations of N2 nematodes at various developmental stages were grown on 10-cm NGM plates (Sulston and Hodgkin, 1988) and washed off with M9 buffer. The nematodes were pelleted and 5× concentrated gel sample buffer (Laemmli, 1970) was added to a final concentration of 1×. The nematodes were boiled for 5 min and the dissolved proteins were fractionated by SDS electrophoresis on a 10% acrylamide/0.27% bis-acrylamide gel (Laemmli, 1970). The proteins were electroblotted onto a nitrocellulose filter by the method of Towbin *et al.* (1979) in 25 mM Tris-HCl, 192 mM glycine buffer, pH 8.3, 20% methanol at 200 mA for 2 h. The filter was blocked in TBS (20 mM Tris-HCl, pH 7.5, 500 mM NaCl)/3% BSA for 30 min and incubated in the primary antibody for 1 h at room temperature in TBS/1% BSA. The filter was washed with TBS for 30 min and then incubated in the secondary antibody for 1 h at room temperature. Goat anti-rabbit IgG (Fc) conjugated to alkaline phosphatase (Promega, Madison, WI) was used as the secondary antibody at a dilution of 1:7500 in TBS/1% BSA. After washing in TBS for 10 min, we developed the blot with nitro blue tetrazolium (NBT) and 5-bromo-4-chloro-3-indolyl phosphate (BCIP) according to the manufacturer's instructions.

Isolation and mapping of cosmid clones

A cosmid library was constructed from N2 (Bristol) genomic DNA that was partially digested with *Xho* II. The DNA was size-fractionated by field inversion gel electrophoresis and subcloned into the *Bam*HI site of pWEKan (J. Mendel, unpublished). The resulting cosmid library contains about three or four genome equivalents, and the average insert size is 35–50 kilobase pairs. The library was gridded in microtiter plates and replicated to nitrocellulose filters. PCR fragments from the *gpa-1, gpa-2,* and *gpa-3* genes were labeled with a Multiprime kit (Amersham) using α-^{32}P-dCTP and hybridized to the nitrocellulose filter replicas in 1% BSA, 7% SDS, 1 mM EDTA, 0.5% formamide, 0.2M $NaPO_4$, pH 7.2 at 65°C for 12 h using 4× 10^5 cpm/ml of probe. The filters were washed in 2× SSC at 50°C for 30 min, 0.2× SSC at 50°C for 30 min, 0.2× SSC at 65°C for 30 min and then exposed to X-ray film. Positive clones were picked, DNA was prepared from them, and their identities confirmed by PCR analysis. The cosmid clones were mapped by Alan Coulson and John Sulston (MRC, Cambridge, England) using a fingerprinting technique that compares the pattern of *Hin*dIII-*Sau*3A1 restriction fragments of a cosmid to that of an overlapping set of cosmids (contigs) that now represents >95% of the *C. elegans* genome (Coulson *et al.,* 1986; A. Coulson, J. Sulston, and R. Waterston, personal communication).

DNA sequence analysis

DNA sequencing reactions were performed using the Sequenase, version 2.0 kit according to the instructions provided by the manufacturer (United States Biochemical, Cleveland, OH). Double stranded miniprep DNA for use as a sequencing template was prepared by a modification of the boiling lysis method of Holmes and Quigley (1981). In this procedure 1.5-ml cultures were grown in Luria broth with ampicillin (100 μg/ml) for 6–8 h or just until the cultures were saturated. The cells were pelleted in 1.5-ml microfuge tubes and resuspended in 300 μl STET (8% sucrose, 5% Triton X-100, 50 mM EDTA, 50 mM Tris-HCl, pH 8.0). Twenty microliters of lysozyme solution (10 mg/ml in 50 mM Tris-HCL, pH 8.0) were added. After incubation at room temperature for 1 min, the mixture was boiled for 2 min and immediately microfuged at room temperature for 5 min. The pellet was removed with a sterile toothpick and 300 μl 75% isopropanol, 2.5 M ammonium acetate was added. After mixing, the tubes were microfuged at room temperature for 5 min. The supernatant was aspirated and the pellet was washed in 1 ml 70% ethanol, 1 ml 100% ethanol, then dried in a Spin-Vac. The pellet was resuspended in 50 μl H_2O. The plasmid DNA was denatured

```
TGTACTTCTAGAACTTCAAGATTCACATTTATATTTTTTTAGTTTTCTATCAAGTCTTGA 2040
                                                sPheLeuSerSerLeuAs

TCGTATCAAGTTACCCGATTATAATCCAACTGAACAGGATATTCTTCTGTCTAGAATCAA 2100
pArgIleLysLeuProAspTyrAsnProThrGluGlnAspIleLeuLeuSerArgIleLy

GACAACTGGAATTGTTGAAGTTAAATTTCAAATGAAAAGTGTAGATTTTAGGTGTGAATA 2160
sThrThrGlyIleValGluValLysPheGlnMetLysSerValAspPheAr

GTTAATGATTCATGAGTTTGCAAGACATGAGTTTCAGGGTATTGGACGTTGGAGGTCAAC 2220
                                            gValPheAspValGlyGlyGlnA

GATCAGAGAGAAAGAAATGGATTCATTGTTTTGAAGATGTTAATGCAATTATTTTCATCG 2280
rgSerGluArgLysLysTrpIleHisCysPheGluAspValAsnAlaIleIlePheIleA

CTGCAATTTCAGAATATGATCAAGTTCTGTTTGAAGATGAGACGACGGTTAGAACAAGGA 2340
laAlaIleSerGluTyrAspGlnValLeuPheGluAspGluThrThr

AAACTTTATGGTTTATTTGAGAATTGTTTTTTTTTTTGCAGAATCGAATGATTGAATCAA 2400
                                                AsnArgMetIleGluSerM

TGAGGCTGTTTGAATCAATTTGTAATTCACGATGGTTCATCAATACTTCAATGATTCTTT 2460
etArgLeuPheGluSerIleCysAsnSerArgTrpPheIleAsnThrSerMetIleLeuP

TCTTAAATAAGAAGGATTTGTTTGCCGAGAAGATTAAAAGAACTTCAATTAAATCCGCAT 2520
heLeuAsnLysLysAspLeuPheAlaGluLysIleLysArgThrSerIleLysSerAlaP

TCCCAGACTATAAAGGTAAGATTTAATTGAACTTTTTCACACTCTGCCCATATATTTTCA 2580
heProAspTyrLysG

GGTGCACAAACCTACGACGAATCGTGTCGATATATTGAGGAAAAGTTTGATGGATTGAAT 2640
 lyAlaGlnThrTyrAspGluSerCysArgTyrIleGluGluLysPheAspGlyLeuAsn

GCAAATCCAGAAAAGACGATTTACATGCATCAGACGTGTGCAACAGATACAGATCAAGTA 2700
AlaAsnProGluLysThrIleTyrMetHisGlnThrCysAlaThrAspThrAspGlnVal

CAAATGATTTTGGACTCAGTGATTGATATGATTATCCAGGCCAACTTGCAAGGATGCGGT 2760
GlnMetIleLeuAspSerValIleAspMetIleIleGlnAlaAsnLeuGlnGlyCysGly

TTGTACTGAAATGTTTGTGCCTTTCTCTCTCTCTCTCTTTCTCTCGCTATTTCTATCACT 2820
LeuTyrEnd

TGGAAATACACATAGGGAAAGCATTTTCAGAAAAAAAAACTGGATTACTTTCTGTTAAGAA 2880
GTAACCTGAAAATATACTTAATCGAAATTTAAGACAAACATGGTAAAGAAGGAAACATGG 2940
GTAATGCCAGGTCAATTCTTTGTTATGA                                  2970
ATGCATTCATTTTGAGTCTATTTTTGTATGCTTCTTTTTTATGTCTAAGCATTCGTCGAC   60
ATTTGAAATATGCGTGAAATAATTTTCCAGTGGAACTAGAGCTGGGGGGCTTCCTATTGA  120
TTTTTGATTGATTTTTGAAAAAGACTTCGACCACGAATATCAATAAAATAATCTAAGTTT  180
CCGGGAAGACTCGAAACATCAAAAAGAAACGACTGGAACTATTCCATCGAACAATAGTCT  240
ATCTTTCGTTCATTTAGATGTGAAATTCATGATGATGTACAAATTGATTCTATTCGAAAG  300
```

Figure 9

```
ATTAATGTATTATTGAACATTTTCAGCTGAATCGAACATCTTTTATCACTCATTATTATG  360
ATATTTGTTCATTTTCAATGCGTCACAGCTTGGTGTCGTTTTGTTGAATTGGCAGCACCA  420
TTTCTTGTGAAAAAATGCTTGCTGTCGTCTTTCATTATTGTCGAGGTGAAACAGAAAAAG  480
AAAGAGACACACGAGATATCAACAACACAAAATTCTGAATATCTGACTTTGGCTCTGACC  540
GCCGTCGCCGCCCGCCCGCTAACGAGCATTTTCTCATTTTCTTCTCATTTACCCCGTGTC  600
TACTTCTTTTCTTATTTTTCGTTTCGTTAAACTTCGGTTGTCGCATAATCAGGCCTGCTT  660
TTCATTTTTCACTTTTTGTTTGAACTTTTGGGCTTTTTAGAGTTGTGCTTCATTCAAATC  720

TATATTCAATTACAGAAAATTATTAGCCAAAGAACAATTATGGGATTATGCCAATCTGCA  780
                                        MetGlyLeuCysGlnSerAla

GAGGACAAAGAACTGACGTTGAAATCAAAAGCGATCGATAAGGAAATGATGCAGAACCAT  840
GluAspLysGluLeuThrLeuLysSerLysAlaIleAspLysGluMetMetGlnAsnHis

ATGTCACAGCAAAAAGTTGTGAAGCTTCTGCTTTTGGGTAAAATTTTAAAACTTATTTTT  900
MetSerGlnGlnLysValValLysLeuLeuLeuLeuG

TGTATTCAGCTTCGCGCAGTCAGAAATTGTTTTACGGTAATCATATGACTCATTGAGAAC  960
TTGTTTGACTTTTTGATAATACGCCGGAATCAAATACAAGCAAAACAAAAGTGTCGTGTG 1020
AGAAATTGTTACCCGCTTTTGTTGAAAAGTCCAAAAAATTTAATTTATGAGAACGTTGGG 1080
ATAAAAAATATTATCGTTTCAGAAGAAATGTTCCCTTCTGTTCTCGAACACTTTTTCAAA 1140
AAATTACTTCAAATTTGTAATAACTGTAATTAATAATTCAGAGAATGAAAGAGCTGTGAT 1200
TTTTAAATACTATATTTACAGACAAAAAGGATAAGTCAGAGAAAATGAGAAATGATCAGT 1260
AATAAGAATAAAGAACATCAACTCAATTTTCTCGAATCATATAAAATAATTATTAAAGTT 1320
TTAAGTTTCTTCTCTCACCCGGCTTCTTGCAAACACTACTTTTTTAACTGCCCGCAAAAA 1380
TTTTCATGTGTTCTAAATATTAATAATCAAATTATTTATCAAGTTATAATTCTCTTTTGA 1440
AGCTTATCAAAGAACTACAAACTACGTAGAACCCGTTGATGTCAAATTTTCAGATATCAA 1500
GTGTGCGTACGCTTAGGTTTTCAGCTGTTTCTTTCATTTGATTCTGAGAATTTAGAAGTT 1560
GAAAAATAAATTAGCTGGGGTTTGTGACAAGTTGGGCAAAATTAAACTCTTACTAATTTC 1620

AGGTGCTGGAGAATGTGGAAAAAGTACGGTTCTGAAACAAATGAGGATTCTTCACGATCA 1680
  lyAlaGlyGluCysGlyLysSerThrValLeuLysGlnMetArgIleLeuHisAspHi

CGGATTTACTGCAGAAGAGGCTGAACAACAGAAAAGTGTCGTTTTCAATAATACCCTTCA 1740
sGlyPheThrAlaGluGluAlaGluGlnGlnLysSerValValPheAsnAsnThrLeuGl

AGCAATGACTGCAATTCTGAAAGGAATGGAAGCACTTCGAATGACCTTTGATAAGCCAAT 1800
nAlaMetThrAlaIleLeuLysGlyMetGluAlaLeuArgMetThrPheAspLysProIl

TCGAGAAAATGATGCAAAATTTGTGATGGAGTCTCATAAAATGCTCCAAGAAGCGAAAGT 1860
eArgGluAsnAspAlaLysPheValMetGluSerHisLysMetLeuGlnGluAlaLysVa

TTTCCCAGAAGAATTAGCAAATGCCATTCAAGCATTATGGAATGATAAAGCCGTTCAGCA 1920
lPheProGluGluLeuAlaAsnAlaIleGlnAlaLeuTrpAsnAspLysAlaValGlnGl

AGTTATTGCAAAAGGAAATGAGTTCCAAATGCCTGAAAGTGCACCACAGTAAGATTTTAC 1980
nValIleAlaLysGlyAsnGluPheGlnMetProGluSerAlaProHi
```

Figure 9. **DNA sequence of *C. elegans gpa-3* gene.** The predicted amino acid sequence of the *gpa-3* protein is shown below the DNA sequence. The standard three-letter amino acid abbreviations are given in the legend to Figure 2. Coding sequences are shown in boldface lettering. The primers used in PCR reactions are underlined. oMP19 corresponds to nucleotide positions 2235–2252, oMP20/oMP21 to positions 2460–2477, and Tα29 to positions 2674–2691.

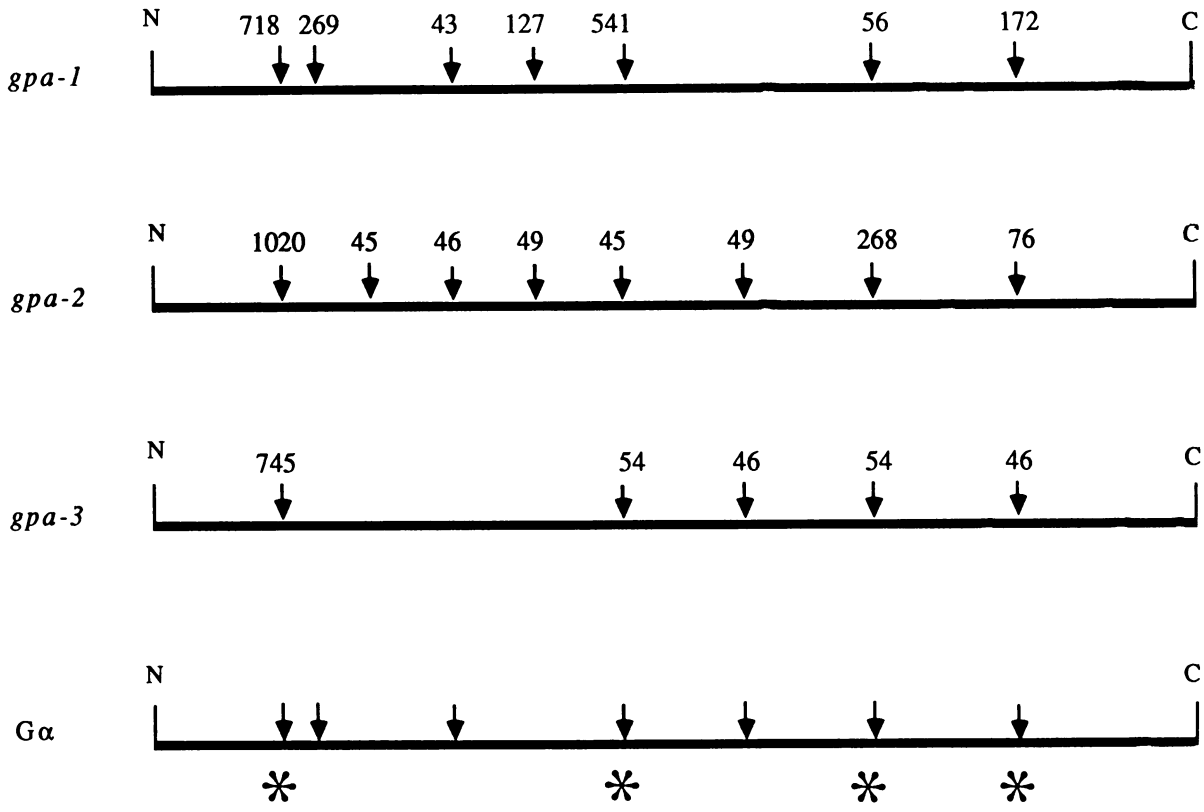

Figure 10. **Intron sizes and locations in *gpa-1*, *gpa-2*, and *gpa-3* coding sequences.** The solid lines represent the coding regions of *gpa-1*, *gpa-2*, *gpa-3*, and Gα (mammalian $G_i\alpha$, $G_o\alpha$, or $T_r\alpha$). The arrows indicate the location of introns and the numbers indicate their size in base pairs. The asterisks indicate intron positions that are conserved in all of the genes. Abbreviations used are N, amino terminus; C, carboxyl terminus.

with alkali by adding 5 μl 2 M NaOH and 2 mM EDTA and incubating at room temperature for 5 min. The solution was neutralized by adding 25 μl 0.9 M sodium acetate, pH 5.3. Plasmid DNA was precipitated by adding 200 μl ethanol, mixing, incubating at −70°C for 5 min, and microfuging for 10 min at 4°C. The supernatant was aspirated and the pellet was washed in 70% ethanol and dried in a Spin-Vac. The denatured DNA was resuspended in 20 μl H_2O. For DNA sequencing reactions, 7 μl denatured template was annealed to 1 μl (10 ng) primer with 2 μl 5× Sequenase buffer at 37°C for 20 min. For DNA sequence analysis of the $G_o\alpha$ cDNA, restriction fragments of λCe6-2 were subcloned into pBluescript KS/−. Most of the sequence of these inserts was determined using sequencing primers that are complementary to pBluescript KS/−. The remaining sequence was determined using synthetic oligonucleotides as insert specific sequencing primers. A T7 DNA polymerase stop obscuring the sequence of positions 171–181 was resolved by the use of the TaqTrack DNA sequencing system (Promega) at a reaction temperature of 80°C. For DNA sequence analysis of the *gpa-1* and *gpa-3* genes, *Hin*dIII restriction fragments of the *gpa-1* and *gpa-3* cosmid clones that hybridize to the *gpa-1* and *gpa-3* PCR-derived probes were identified by Southern blotting and subcloned into pBluescript KS/− (see Figures 6A and 8A). To obtain the complete sequence of the *gpa-3* gene, it was also necessary to subclone a *Nsi* I restriction fragment that overlaps the *Hin*dIII fragment. The DNA sequence of the *gpa-1* and *gpa-3* genes was determined by a combination of three methods. First, specific restriction fragments of the original subclones were subcloned further and sequenced as described above. Second, γδ transposons were inserted at random sites in subclones constructed in pMOB (Strathmann *et al.*, 1991). The location of the γδ transposons was determined by PCR, and DNA sequences were determined by the use of transposon-specific primers corresponding to each of the different ends of γδ. Because γδ generates a five base pair duplication at the site of insertion, it is possible to obtain sequences that overlap at the site of γδ insertion. Third, any remaining sequences that were not obtained by the first two methods were determined by constructing synthetic primers based on available sequence and using these as sequencing primers. Sequence compressions were resolved using dITP instead of dGTP in the sequencing reactions; T7 polymerase stops were resolved by the use of the TaqTrack system at a reaction temperature of 80°C. DNA sequences were analyzed using Pustell programs (Pustell and Kafatos, 1984) run on an IBM PC or the University of Wisconsin Genetics Computer Group sequence analysis software package (Devereux *et al.*, 1984) run on a VAXstation 2000.

Acknowledgments

We thank Mike Strathmann for advice on PCR reactions and for the provision of materials for the γδ transposon-

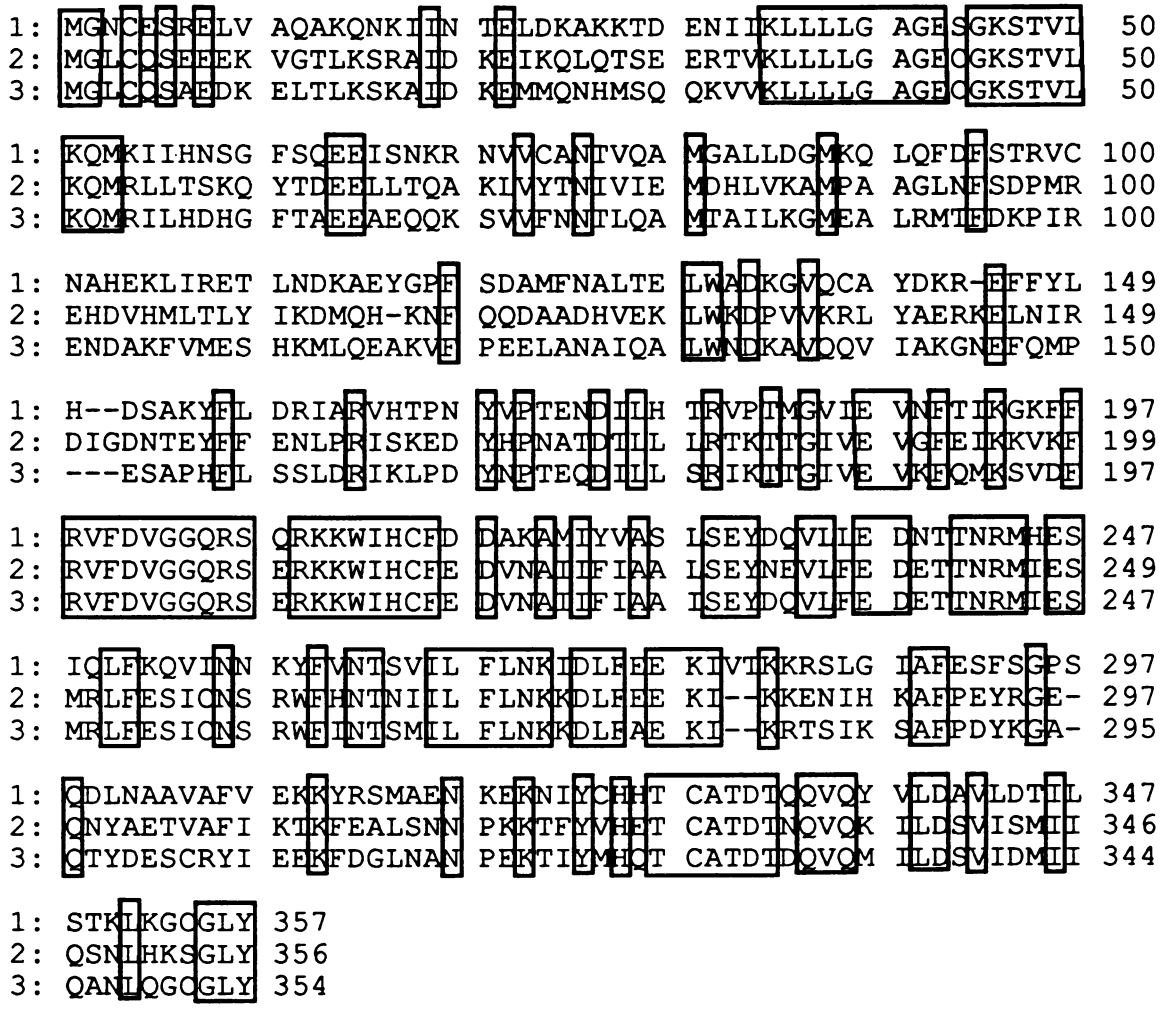

Figure 11. **Comparison of the amino acid sequences of *C. elegans gpa-1, gpa-2,* and *gpa-3.*** Sequences were aligned using the FASTP program of Lipman and Pearson (1985). Dashes indicate gaps introduced to obtain an optimal alignment. One-letter amino acid abbreviations are given in the legend to Figure 3. Amino acids that are found in all three sequences at the same position are boxed.

based sequencing method, the Caltech Microchemical Facility for oligonucleotide synthesis, Andrea Holboke for *C. elegans* genomic DNA, Alan Coulson and John Sulston for gene mapping data, Eva Neer and Graeme Milligan for $G_o\alpha$ antibodies, Stuart Kim and Russell Hill for the *C. elegans* cDNA library, and Yasumi Ohshima and Ron Plasterk for generous communication of results before publication. P.W.S. is an investigator of the Howard Hughes Medical Institute. Supported by a Developmental Biology Grant from the Lucille P. Markey Foundation and by USPHS grants GM34236 to M.I.S. and HD23690 to P.W.S.

Received: October 5, 1990.
Revised and accepted: November 16, 1990.

References

Ballinger, D.G., and Benzer, S. (1989). Targeted gene mutations in *Drosophila.* Proc. Natl. Acad. Sci. USA *86,* 9402–9406.

Breathnach, R., and Chambon, P. (1981). Organization and expression of eucaryotic split genes coding for proteins. Annu. Rev. Biochem. *50,* 349–383.

Brenner, S. (1974). The genetics of *Caenorhabditis elegans.* Genetics *77,* 71–94

Buss, J.E., Mumby, S., Casey, P.J., Gilman, A.G. and Sefton, B. (1987). Myristoylated α subunits of guanine nucleotide-binding regulatory proteins. Proc. Natl. Acad. Sci. USA *84,* 7493–7497.

Caron, M.G. (1989). The guanine nucleotide regulatory protein-coupled receptors for nucleosides, nucleotides, amino acids, and amine neurotransmitters. Curr. Opinion Cell Biol. *1,* 159–166.

Casey, P., Fong, H.K.W., Simon, M.I. and Gilman, A.G. (1990). G_z, a guanine nucleotide-binding protein with unique biochemical properties. J. Biol. Chem. *265,* 2383–2390.

Cerione, R.A., Regan, J.W., Nakata, H., Codina, J., Benovic, J.L., Gierschik, P., Somers, R., Spiegel, A.M., Birnbaumer,

L., Lefkowitz, R.J., and Caron, M.G. (1986). Functional reconstitution of the α_2-adrenergic receptor with guanine nucleotide regulatory proteins in phospholipid vesicles. J. Biol. Chem. *261,* 3901–3909.

Chalfie, M., and White, J. (1988). The nervous system. In: The Nematode *Caenorhabditis elegans,* ed. W.B. Wood, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 337–391.

Coulson, A., Sulston, J., Brenner, S., and Karn, J. (1986). Toward a physical map of the genome of the nematode *Caenorhabditis elegans.* Proc. Natl. Acad. Sci. USA *83,* 7821–7825.

Coulson, A., Waterson, R., Kiff, J., Sulston, J., and Kohara, Y. (1988). Genome linking with yeast artificial chromosomes. Nature *335,* 184–186.

deSousa, S.M., Hoveland, L.L., Yarfitz, S., and Hurley, J.B. (1989). The *Drosophila* $G_o\alpha$-like G protein gene produces multiple transcripts and is expressed in the nervous system and in ovaries. J. Biol. Chem. *264,* 18544–18551.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. *12,* 387–395.

Emmons, S.W. (1988). The genome. In: The Nematode *Caenorhabditis elegans,* ed. W.B. Wood, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 47–80.

Emmons, S.W., Rosenzweig, B. and Hirsh, D. (1980). Arrangement of repeated sequences in the DNA of the nematode *Caenorhabditis elegans.* J. Mol. Biol. *144,* 481–500.

Feinberg, A.P. and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction fragments to high specific activity. Anal. Biochem. *132,* 6–13.

Fino Silva, I., and Plasterk, R.H.A. (1990). Characterization of a G-protein α subunit gene from the nematode *Caenorhabditis elegans.* J. Mol. Biol. *215,* 483–487.

Fire, A. (1986). Integrative transformation of *Caenorhabditis elegans.* EMBO J. *5,* 2673–2680.

Freissmuth, M., and Gilman, A.G. (1989). Mutations of $G_{s\alpha}$ designed to alter the reactivity of the protein with bacterial toxins. Substitutions at arg^{187} result in loss of GTPase activity. J. Biol. Chem. *264,* 21907–21914.

Gabrion, J., Brabet, P., Dao, B.N.T., Homburger, V., Dumuis, A., Sebben, M., Rouot, B., and Bockaert, J. (1989). Ultrastructural localization of the GTP-binding protein G_o in neurons. Cell Signaling *1,* 107–123.

Gillespie, P.G., and Beavo, J.A. (1988). Characterization of a bovine cone photoreceptor phosphodiesterase purified by cyclic GMP-sepharose chromatography. J. Biol. Chem. *263,* 8133–8141.

Gilman, A.G. (1987) G proteins: transducers of receptor-generated signals. Annu. Rev. Biochem. *56,* 615–650.

Graziano, M.P. and Gilman, A.G. (1989). Synthesis in *Escherichia coli* of GTPase-deficient mutants of $G_s\alpha$. J. Biol. Chem. *264,* 15475–15482.

Gunderson, R.E., and Devreotes, P.N. (1990). In vivo receptor-mediated phosphorylation of a G protein in *Dictyostelium.* Science *248,* 591–593.

Harris-Warrick, R.M., Hammond, C., Paupardin-Tritsch, D., Homburger, V., Rouot, B., Bockaert, J., and Gerschenfeld, H.M. (1988). An α_{40} subunit of a GTP-binding protein immunologically related to G_o mediates a dopamine-induced decrease of Ca^{2+} current in snail neurons. Neuron *1,* 27–32.

Hescheler, J., Rosenthal, W., Trautwein, W., and Schultz, G. (1987). The GTP-binding protein, G_o, regulates neuronal calcium channels. Nature *325,* 445–447.

Hodgkin, J., Edgley, M., Riddle, D.L., and Albertson, D.G. (1988). Genetic map. In: The Nematode *Caenorhabditis elegans,* ed. W.B. Wood, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 491–583.

Hodgkin, J., Horvitz, H.R., and Brenner, S. (1979). Non-disjunction mutants of the nematode *Caenorhabditis elegans.* Genetics *91,* 67–94.

Holmes, D.S., and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. Anal. Biochem. *114,* 193–197.

Hsu, W.H., Rudolph, U., Sanford, J., Bertrand, P., Olate, J., Nelson, C., Moss, L.G., Boyd, A.E., III, Codina, J., and Birnbaumer, L. (1990). Molecular cloning of a novel splice variant of the α subunit of the mammalian G_o protein. J. Biol. Chem. *265,* 11220–11226.

Huang, X.-Y., and Hirsh, D. (1989). A second trans-spliced RNA leader sequence in the nematode *Caenorhabditis elegans.* Proc. Natl. Acad. Sci. USA *86,* 8640–8644.

Itoh, H., Toyama, R., Kozasa, T., Tsukamoto, T., Matsuoka, M., and Kaziro, Y. (1988). Presence of three distinct molecular species of G_i protein α subunit. Structure of rat cDNAs and human genomic DNAs. J. Biol. Chem. *263,* 6656–6664.

Jones, D.T., and Reed, R.R. (1987). Molecular cloning of five GTP-binding protein cDNA species from rat olfactory neuroepithelium. J. Biol. Chem. *262,* 14241–14249.

Jones, T.L.Z., Simonds, W.F., Merendino, J.J., Jr., Brann, M.R., and Spiegel, A.M. (1990). Myristoylation of an inhibitory GTP-binding protein α subunit is essential for its membrane attachment. Proc. Natl. Acad. Sci. USA *87,* 568–572.

Kahano, Y., Tsai, S.-C., Adamik, R., Hewlett, E.L., Moss, J., and Vaughan, M. (1984). Rhodopsin-enhanced GTPase activity of the inhibitory GTP-binding protein of adenylate cyclase. J. Biol. Chem. *259,* 7378–7381.

Kaiser, K., and Goodwin, S.F. (1990). "Site-selected" transposon mutagenesis of *Drosophila.* Proc. Natl. Acad. Sci. USA *87,* 1686–1690.

Kaziro, Y., Itoh, H., Kozasa, T., Toyama, R., Tsukamoto, T., Matsuoka, M., Nakafuku, M., Obara, T., Takagi, T., and Hernandez, R. (1988). Structures of the genes coding for G-protein α subunits from mammalian and yeast cells. Cold Spring Harbor Symp. Quant. Biol. *53,* 209–220.

Kenyon, C. (1988). The nematode *Caenorhabditis elegans.* Science *240,* 1448–1453.

Krause, M., and Hirsh, D. (1987). A *trans*-spliced leader sequence on actin mRNA in *C. elegans.* Cell *49,* 753–761.

Laemmli, U. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature *227,* 680–685.

Landis, C.A., Masters, S.B., Spada, A.M., Bourne, H.R., and Vallar, L. (1989). GTPase inhibiting mutations activate the α chain of G_s and stimulate adenylate cyclase in human pituitary tumours. Nature *340,* 692–696.

Lavu, S., Clark, J., Swarup, R., Matsushima, K., Paturu, K., Moss, J., and Kung, H.-F. (1988). Molecular cloning and DNA sequence analysis of the human guanine nucleotide binding

protein $G_o\alpha$. Biochem. Biophys. Res. Commun. *150,* 811–815 and erratum *153,* 487.

Lipman, D.J., and Pearson, W.R. (1985). Rapid and sensitive protein similarity searches. Science *227,* 1435–1441.

Lochrie, M.A., and Simon, M.I. (1988). G protein multiplicity in eukaryotic organisms. Biochemistry *27,* 4957–4965.

Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory.

Masters, S.B., Miller, R.T., Chi, M.-H., Chang, F.-H., Beiderman, B., Lopez, N.G., and Bourne, H.R. (1989). Mutations in the GTP-binding site of $G_s\alpha$ alter stimulation of adenylyl cyclase. J. Biol. Chem. *264,* 15467–15474.

McFadzean, I., Mullaney, I., Brown, D.A., and Milligan, G. (1989). Antibodies to the GTP binding protein, G_o, antagonize noradrenaline-induced calcium current inhibition in NG108-15 hybrid cells. Neuron *3,* 177–182.

Milburn, M.V., Tong, L., deVos, A.M., Brunger, A., Yamaizumi, Z., Nishimura, S., and Kim, S.-H. (1990). Molecular switch for signal transduction: structural differences between active and inactive forms of the protooncogenic ras proteins. Science *247,* 939–945.

Moriarty, T.M., Padrell, E., Carty, D.J., Omri, G., Landau, E.M., and Iyengar, R. (1989). G_o protein as signal transducer in the pertussis toxin-sensitive phosphatidylinositiol pathway. Nature *343,* 79–82.

Moss, J., and Vaughan, M. (1988). ADP-ribosylation of guanyl nucleotide-binding regulatory proteins by bacterial toxins. Adv. Enzymol. *61,* 303–379.

Mumby, S.M., Heukeroth, R.O., Gordon, J.I., and Gilman, A.G. (1990). G-protein α-subunit expression, myristoylation, and membrane association in COS cells. Proc. Natl. Acad. Sci. USA *87,* 728–732.

Neer, E.J., Lok, J.M., and Wolf, L.G. (1984). Purification and properties of the inhibitory guanine nucleotide regulatory unit of brain adenylate cyclase. J. Biol. Chem. *259,* 14222–14229.

Olate, J., Jorquera, H., Purcell, P., Codina, J., Birnbaumer, L., and Allende, J.E. (1989). Molecular cloning and sequence determination of a cDNA coding for the α-subunit of a G_o-type protein of *Xenopus laevis* oocytes. FEBS Lett. *244,* 188–192.

Proudfoot, N.J., and Brownlee, G.G. (1976). 3′ Non-coding region sequences in eukaryotic messenger RNA. Nature *263,* 211–214.

Provost, N.M., Somers, D.E., and Hurley, J.B. (1988). A *Drosophila melanogaster* G protein α subunit gene is expressed primarily in embryos and pupae. J. Biol. Chem. *263,* 12070–12076.

Pustell, J., and Kafatos, F.C. (1984). A convenient and adaptable package of computer programs for DNA and protein sequence management, analysis, and homology determination. Nucleic Acids Res. *12,* 643–655.

Raport, C.J., Dere, B., and Hurley, J.B. (1989). Characterization of the mouse rod transducin α subunit gene. J. Biol. Chem. *264,* 7122–7128.

Sagi-Eisenberg, R. (1989). GTP-binding proteins as possible targets for protein kinase C action. Trends Biochem. Sci. *14,* 355–357.

Schmidt, C.J., Garen-Fazio, S., Chow, Y.-K., and Neer, E. (1989). Neuronal expression of a newly identified *Drosophila melanogaster* G protein α_o subunit. Cell Regul. *1,* 125–134.

Senogles, S.E., Spiegel, A.M., Padrell, E., Iyengar, R., and Caron, M. (1990). Specificity of receptor G-protein interactions. Discrimination of G_i subtypes by the D_2 dopamine receptor in a reconstituted system. J. Biol. Chem. *265,* 4507–4514.

Serventi, I.M., Price, S.R., Murtagh, J.J., and Moss, J. (1990). Structure and function of $G_o\alpha$. In: ADP-Ribosylating Toxins and G Proteins, ed. J. Moss and M. Vaughan, Washington, DC: American Society for Microbiology, 325–348.

Sternweis, P.C., and Robishaw, J.D. (1984). Isolation of two proteins with high affinity for guanine nucleotides from membranes of bovine brain. J. Biol. Chem. *259,* 13806–13813.

Strathmann, M., Hamilton, B.A., Mayeda, C., Simon, M.I., Meyerowitz, E.M., and Palazzolo, M.J. (1991). Transposon-facilitated DNA sequencing. Proc. Natl. Acad. Sci. USA (in press).

Strathmann, M., and Simon, M.I. (1990). G protein diversity: a distinct class of alpha subunits is present in vertebrates and invertebrates. Proc. Natl. Acad. Sci. USA *87,* 9113–9117.

Strathmann, M.P., Wilke, T.M., and Simon, M.I. (1989). Diversity of the G-protein family: sequences from five additional α subunits in mouse. Proc. Natl. Acad. Sci. USA *86,* 7407–7409.

Strathmann, M., Wilke, T.M., and Simon, M.I. (1990). Alternative splicing produces transcripts encoding two forms of the α subunit of GTP-binding protein G_o. Proc. Natl. Acad. Sci. USA *87,* 6477–6481.

Strittmatter, S.M., Valenzuela, D., Kennedy, T.E., Neer, E., and Fishman, M. (1990). G_o is a major protein of growth cone protein subject to regulation by GAP-43. Nature *344,* 836–841.

Sulston, J. (1988) Cell lineage. In: The Nematode *Caenorhabditis elegans,* ed. W.B. Wood, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 123–155.

Sulston, J., and Hodgkin, J. (1988). Methods. In: The Nematode *Caenorhabditis elegans,* ed. W.B. Wood. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 587–606.

Sulston, J., Horvitz, H.R., and Kimble, J. (1988). Cell lineage. In: The Nematode *Caenorhabditis elegans,* ed. W.B. Wood, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 457–489.

Thambi, N.C., Quan, F., Wolfgang, W.J., Spiegel, A., and Forte, M. (1989). Immunological and molecular characterization of $G_o\alpha$-like proteins in the *Drosophila* central nervous system. J. Biol. Chem. *264,* 18552–18560.

Towbin, H., Staehlin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc. Natl. Acad. Sci. USA *76,* 4350–4354.

Towler, D.A., Gordon, J.I., Adams, S.P., Glaser, L. (1988). The biology and enzymology of eukaryotic protein acylation. Annu. Rev. Biochem. *57,* 69–99.

Ueda, H., Uno, S., Harada, J., Kobayashi, I., Katada, T., Ui, M., and Satoh, M. (1990). Evidence for receptor-mediated inhibition of intrinsic activity of GTP-binding protein, G_i1 and

G_i2, but not G_o in reconstitution experiments. FEBS Lett. *266,* 178–182.

Van der Voorn, L., Gebbink, M., Plasterk, R.H.A., and Ploegh, H.L. (1990). Characterization of a G-protein β-subunit gene from the nematode *Caenorhabditis elegans.* J. Mol. Biol. *213,* 17–26.

Van Dongen, A.M.J., Codina, J., Olate, J., Mattera, R., Joho, R., Birnbaumer, L., and Brown, A.M. (1988). Newly identified brain potassium channels gated by the guanine nucleotide binding protein G_o. Science *242,* 1433–1437.

Van Meurs, K.P., Angus, C.W., Lavu, S., Kung, H.-F., Czarnecki, S.K., Moss, J., and Vaughan, M. (1987). Deduced amino acid sequence of bovine retinal $G_o\alpha$: homologies to other guanine nucleotide-binding proteins. Proc. Natl. Acad. Sci. USA *84,* 3107–3111.

Weinstein, L.S., Spiegel, A.M., and Carter, A.D. (1988). Cloning and characterization of the human gene for the α-subunit of G_i2, a GTP-binding signal transduction protein. FEBS Lett. *232,* 333–340.

White, J., Southgate, E., and Durbin, R. (1988). Neuroanatomy. In: The Nematode *Caenorhabditis elegans,* ed. W.B. Wood, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 433–455.

White, J., Southgate, E., Thomson, J.N., and Brenner, S. (1986). The structure of the nervous system of the nematode *Caenorhabditis elegans.* Philos. Trans. R. Soc. Lond. Biol. Sci. *314,* 1–340.

Wickens, M., and Stephenson, P. (1984). Role of the conserved AAUAAA sequence: Four AAUAAA point mutations prevent messenger RNA 3′ end formation. Science *226,* 1045–1051.

Wood, W.B. (1988). Introduction to *Caenorhabditis elegans* biology. In: The Nematode *Caenorhabditis elegans,* ed. W.B. Wood, Cold Spring Harbor, New York: Cold Spring Harbor Laboratory, 1–16.

Worley, P.F., Baraban, J. M., Van Dop, C., Neer, E.J., and Snyder, S.H. (1986). G_o, a guanine nucleotide-binding protein: immunohistochemical localization in rat brain resembles distribution of second messenger systems. Proc. Natl. Acad. Sci. USA *83,* 4561–4565.

Yatani, A., Mattera, R., Codina, J., Graf, R., Okabe, K., Padrell, E., Iyengar, R., Brown, A., and Birnbaumer, L. (1988). The G protein-gated atrial K^+ channel is stimulated by three distinct $G_i\alpha$-subunits. Nature *336,* 680–682.

Yoon, J., Shortridge, R.D., Bloomquist, B.T., Schneuwly, S., Perdew, M.H., and Pak, W.L. (1989). Molecular characterization of *Drosophila* gene encoding $G_o\alpha$ subunit homologue. J. Biol. Chem. *264,* 18536–18543.